

**ESSAYS ON THE WORKINGS
AND USES OF FUTURES MARKETS**

A Dissertation

by

HENRY L. BRYANT IV

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2003

Major Subject: Agricultural Economics

**ESSAYS ON THE WORKINGS
AND USES OF FUTURES MARKETS**

A Dissertation

by

HENRY L. BRYANT IV

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

David A. Bessler
(Co-Chair of Committee)

Michael S. Haigh
(Co-Chair of Committee)

Richard Woodward
(Member)

R. Brian Balyeat
(Member)

A. Gene Nelson
(Head of Department)

August 2003

Major Subject: Agricultural Economics

ABSTRACT

Essays on the Workings and Uses of Futures Markets.

(August 2003)

Henry L Bryant IV, B.S., University of Nevada, Las Vegas

Co-Chairs of Advisory Committee: Dr. David A. Bessler
Dr. Michael S. Haigh

This dissertation investigates various issues of interest regarding the workings and uses of commodity futures markets. Chapter II evaluates the relative performances of various estimators of bid-ask spreads in futures markets using commonly available transaction data. Results indicate a wide divergence in the performance of the competing estimators. This chapter also examines the effect of automating trading on spreads in commodity futures markets. Results indicate that spreads generally widened after trading was automated on the markets considered, and the tendency for spreads to widen during periods of high volatility increased. These results are in contrast to those found in higher volume financial futures markets.

Chapter III investigates various unresolved issues regarding futures markets, using formal methods appropriate for inferring causal relationships from observational data when some relevant quantities are hidden. I find no evidence supporting the generalized version of Keynes's theory of normal backwardation. I find no evidence supporting theories that predict that the level of activity of speculators or uninformed traders affects the level of price volatility, either positively or negatively. My evidence strongly supports the mixture of distribution hypothesis (MDH) that trading volume and

price volatility have one or more latent common causes, resulting in their positive correlation.

Chapter IV examines partial equilibrium and statistical approaches to hedging. Different types of hedgers have traditionally used each of two approaches: derivatives dealers and market makers have typically used the former approach to hedge their portfolios, while commodity producers and consumers more commonly use the latter. This research provides the first known comparison of the out-of-sample hedging performance of the two approaches. Results indicate that for a simple derivative with a linear payoff function (a futures contract), the statistical models significantly outperform the partial equilibrium models considered here.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to, Drs. Haigh, Bessler, and Davis. They have taught, motivated, and inspired me. Through their guidance and molding I have become a scholar. Thanks also to the other members of my committee.

Thanks are due to the College of Agriculture and Life Sciences for financial support in the form of the Tom Slick Senior Graduate Research Fellowship, and to the Department of Agricultural Economics for financial support in the form of Robert G. Cherry Fellowships. Thanks also to Dr. Haigh and Dr. Alan Love for financial support in the form of research assistantships.

Thanks to all of the other teachers who have ever taught me. Finally, I deeply appreciate the support and patience of my family as I have made my way through graduate school.

TABLE OF CONTENTS

	Page
ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF FIGURES.....	viii
LIST OF TABLES.....	x
 CHAPTER	
I INTRODUCTION.....	1
II BID-ASK SPREADS IN COMMODITY FUTURES MARKETS.	6
Introduction.....	6
Spread Estimator Performance.....	10
Spreads in Electronic and Open Outcry Commodity Futures Markets.....	26
Conclusions.....	36
III CAUSALITY IN FUTURES MARKETS.....	38
Introduction.....	38
Issues Investigated.....	40
Inferring Causal Relationships From Incomplete, Observational Data.....	48
Data.....	55
Analysis.....	58
Conclusions.....	64
IV COMPARING THE PERFORMANCES OF THE PARTIAL EQUILIBRIUM AND TIME SERIES APPROACHES TO HEDGING.....	66

CHAPTER	Page
Introduction.....	66
Hedging Commodity Price Risk Using Time Series Models.....	69
Hedging Commodity Price Risk Using Partial Equilibrium Models.....	73
Data, Parameter Estimation, and Parameter Inference.....	82
Hedging Effectiveness.....	91
Conclusions.....	96
V CONCLUSIONS.....	98
BIBLIOGRAPHY	101
APPENDIX A FIGURES.....	112
APPENDIX B TABLES.....	129
VITA.....	143

LIST OF FIGURES

FIGURE	Page
2.1 Daily average bid-ask spread for November 2000 coffee futures (dollars per tonne).....	112
3.1 Directed acyclic graph G.....	113
3.2 Inducing path graph G' over $O = \{V_1, V_2, V_4, V_6\}$ associated with directed acyclic graph G.....	114
3.3 Directed acyclic graph H.....	115
3.4 Partially-oriented inducing path graph for Corn.....	116
3.5 Partially-oriented inducing path graph for Crude Oil.....	117
3.6 Partially-oriented inducing path graph for Eurodollars.....	118
3.7 Partially-oriented inducing path graph for Gold.....	119
3.8 Partially-oriented inducing path graph for Japanese Yen.....	120
3.9 Partially-oriented inducing path graph for Coffee.....	121
3.10 Partially-oriented inducing path graph for Live Cattle.....	122
3.11 Partially-oriented inducing path graph for S&P 500.....	123
4.1 The term structure of volatility (TSV) of crude oil in the Gibson-Schwartz model using different parameter estimation techniques.....	124
4.2 The term structure of volatility (TSV) of crude oil in the Schwartz 1997 model using different parameter estimation and inference techniques.....	125
4.3 The option-implied term structure of volatility (TSV) of crude oil in the Schwartz 1997 model observed on two dates.....	126
4.4 Partial time paths of the in-sample optimal hedge ratios generated by the VEC-GARCH and Gibson-Schwartz (GS) models.....	127

FIGURE

Page

4.5	Partial time paths of the in-sample optimal hedge ratios generated by the Schwartz 1997 (S97) models.....	128
-----	--	-----

LIST OF TABLES

TABLE	Page
2.1 Example of LIFFE coffee futures data.....	129
2.2 Deviations of intra-day average spreads from overall daily average spreads.....	130
2.3 Correlations of daily average spreads and estimates of daily average spreads.....	131
2.4 Performance of estimators by commodity.....	132
2.5 Coefficient estimates and p-value for differences in bias and variance components for each pair of bid-ask spread estimators.....	133
2.6 P-values for encompassing tests.....	134
2.7 Determinants of daily average spreads regression results.....	135
3.1 Results of augmented Dickey-Fuller tests.....	136
3.2 Conditioning sets that result in vanishing correlations.....	137
4.1 Results from augmented Dickey-Fuller tests on price data.....	138
4.2 Parameter estimates and residual diagnostics for the univariate GARCH(1,1) models.....	139
4.3 Parameter estimates and residual diagnostics for the multivariate GARCH(1,1) model.....	140
4.4 In-sample hedging effectiveness.....	141
4.5 Out-of-sample hedging effectiveness.....	142

CHAPTER I

INTRODUCTION

Risk is not a “good” in the conventional sense, but it is nonetheless a feature of the economic world with which agents must contend. As is the case with most goods, the initial allocation of risk that is dealt by nature is not likely to be optimal. Some agents in the economy are endowed with an abundance of risk, while others are endowed with relatively little. Significant gains from trade are generally available. Futures contracts are important tools that facilitate the efficient reallocation of risk among agents in an economy. Their importance is demonstrated by the phenomenal growth of futures trading since its beginnings in the mid-nineteenth century. In the United States, the volume of trade in futures contracts is now hundreds of millions of contracts per year, and an impressive volume of trade occurs elsewhere in the world as well. Given the important roll that futures markets plays in the economy, a sound understanding of their workings and uses is critical. This dissertation makes contributions towards such understanding, in three areas.

In Chapter II, I investigate issues regarding bid-ask spreads in commodity futures markets. The bid-ask spread is the difference between the prices available for immediate purchase immediate sale of a futures contract. It is an important source of transaction cost for futures market participants, and has thus received significant attention from the academy in recent years. Much of this attention has been directed towards the study of

This dissertation flows the style and format of the Journal of Futures Markets.

bid-ask spreads in financial futures markets, however, and relatively little attention has been paid to spreads in commodity futures markets. The volume of trade in the large financial futures markets dwarfs that of most commodity futures markets, and it cannot be taken as certain that results found in the former markets apply to the latter. This first line of inquiry thus helps to fill this gap in the extant literature. Two particular issues are investigated: the estimation of spreads in commodity futures markets when they cannot be observed, and the effect of automating trading on bid-ask spreads in these markets.

In the open-outcry trading that is traditional in futures markets, bid ask spreads are generally not observed, and can only be inferred from transaction data. Several spread estimators have thus been proposed in the literature. Evaluation of the performance of these estimators is made difficult, however, by the very fact that spreads are not typically observed. I make use of a unique data set from the London International Financial Futures Exchange (LIFFE) that does include observations of the spread components, as well as the commonly available transaction price data. Use of this LIFFE data thus facilitates an evaluation of the accuracy of spread estimates that might be computed when bid and ask data are not reported, as is the case in the large U.S. commodity futures markets. Traders, regulators, and market microstructure researchers need accurate estimates of bid-ask spreads.

The second issue regarding bid-ask spreads that I investigate is the effect of automating trading. It has been argued that electronic trading should be more efficient than other forms of trading, resulting in lower transaction cost for market participants, and many futures exchanges around the world have moved in this direction, either

partially or fully. The advisability of the remaining open outcry futures markets moving to electronic trading remains the topic of intense debate, however, as some argue that the anonymity of such a system could result in increased rather than decreased transaction costs (due to wider bid-ask spreads). Given this interest, it is not surprising that several studies have investigated the relative magnitudes of spreads in electronic and open outcry settings. Again, however, this previous work has focused on financial rather than commodity futures markets. A further contribution of Chapter II is to evaluate the impact on nominal spread magnitudes of moving from open outcry to electronic trading in two LIFFE commodity futures markets, after controlling for spread determinants. The findings of this research have important implications for market participants and other exchanges that may be contemplating automating trading in their commodity futures markets.

In Chapter III, I investigate the empirical evidence supporting various theories on the working of futures markets. Two difficulties present themselves to the researcher conducting such an investigation. First, relevant quantities are frequently not observed. Second, it is not feasible for researchers to uncover answers through experimentation. Systematic manipulation of futures markets is not only impractical; in many cases it is illegal. These difficulties are not unique to the empirical investigation of futures markets - researchers in numerous fields find themselves operating under such circumstances. This situation has inspired a recent multidisciplinary effort to develop a body of theory concerning the inference of causal relationships using observational data. A subset of this literature further concerns itself with conducting this inference when the

observational data are incomplete. I use these methods to investigate three issues: i) Keynes's theory predicting that futures prices do not equal expected future cash prices in order that hedgers might compensate speculators for providing risk-bearing services, ii) theories explaining the positive correlation between the volume of trade and price volatility that are frequently observed in futures markets, and iii) theories predicting that the activities of some types of traders impact levels of price volatility. Clearly, establishing the truth of Keynes' theory is important for market participants. Should a hedger anticipate losing money on average in exchange for enjoying reduced risk? Can speculators expect to be profitable on average by merely taking appropriate positions considering the needs of hedgers? The second and third issues are important for market participants, researchers, and regulators. All market participants are obviously impacted by price volatility and market depth, and clearly should be interested in the underlying causes. Researchers will be interested in the correct specification of empirical models, and results inconsistent with existing theories might inspire new ones. Regulators have displayed an interest in curbing excessive levels of price volatility, and the success of such an endeavor would be greatly aided by a clear understanding of its causes.

In Chapter IV, I turn from investigating the workings of futures markets to investigating their use. A large existing literature has proposed various schemes for optimally hedging price risk. These schemes come in two varieties: those that make use of statistical time series models, and those that make use of models of the equilibrium price relationships between the underlying asset and the hedging instrument. The former variety has been cultivated in the applied economics literature, while the latter has been

developed in the financial economics literature. The time series approach has generally been applied by those hedging commodity price risk, while the partial equilibrium approach has typically been employed by holders of large derivative portfolios. Either type of hedger might use either approach, however, despite the traditional divide. Thus far, no direct evaluation of the relative performances of the two approaches has been presented in the literature. I carry out such an evaluation, for a hypothetical hedger with a non-tradable long position in a commodity. The importance of this issue is obvious – it directly investigates the best means by which hedgers can increase their utility.

Each of Chapters II through IV is self-contained, each with its own introduction, body and conclusions. A summary of the conclusions is presented in Chapter V, which is followed by bibliographic information, figures and tables.

CHAPTER II

BID-ASK SPREADS IN COMMODITY FUTURES MARKETS

INTRODUCTION

The bid-ask spread, the difference between the price that must be paid for immediate purchase and the price that can be received for immediate sale of a security, is an important source of transaction cost for market participants. It has thus been a primary concern in market microstructure research and has received much attention in recent years. Researchers have investigated such topics as the magnitudes and determinants of bid-ask spreads, the impacts of different market microstructures on spreads, intra-day variations in spreads, and estimating spreads when they cannot be observed. These issues have been studied for equities, debt instruments, futures and options. In the futures markets, research regarding bid-ask spreads has concentrated primarily on the financial markets. In this chapter, we focus on commodity futures markets because unlike financial markets, microstructure issues have not been analyzed in any complete manner and moreover, there is strong reason to believe that some findings from financial markets may not be directly applicable to commodity markets. Consequently, we investigate two issues of recent interest and controversy regarding bid-ask spreads in commodity futures markets.

The first of these issues is the estimation of bid-ask spreads. Bid-ask spreads are often not observed, particularly in open outcry futures markets, necessitating their estimation using transaction data. Accurate estimates of spreads are needed by traders,

regulators, and market microstructure researchers, among others. Several estimators have thus been proposed and implemented in various markets to estimate nominal and effective spreads.¹ Directly evaluating the performance of these estimators is made difficult, however, by the very fact that spreads are not typically observed. Direct evaluations have been carried out, however, in Locke and Venkatesh (1997) and ap Gwilym and Thomas (in press). These studies both suggest that spreads estimators perform poorly in estimating *effective* spreads. However to date there has been no direct evaluation of estimator performance in estimating *nominal* spreads in commodity futures markets. In commodity futures markets, there is a higher proportion of information traders than there tends to be in financial markets (Foster and Viswanathan, 1994). This feature is likely to affect estimator accuracy and so it is not clear that results from financial markets can be immediately applied to the commodity markets. Therefore, given this difference in the markets we apply our bid-ask spread estimators to commodity transaction data and assess their performance in estimating nominal spreads.

¹ Research into spreads can be classified as concerning nominal, effective, or quoted spreads. In this chapter, we define *quoted* spreads as those determined by the bids and offers that officially designated market makers are required to simultaneously quote in a specialist system, such as that of the New York Stock Exchange. In futures markets, there are no officially designated monopolistic market makers, and hence no quoted spreads. Instead, there are simply prevailing best bidding and asking prices (which may be provided by different traders) that together imply a *nominal* spread. Here we define *effective* spreads to be the average transfer of wealth from market participants to liquidity providers. These differ from quoted spreads due to trading inside the quoted prices in a specialist system (Roll 1984; Petersen and Fiakowski, 1994). In futures markets, effective spreads differ from nominal spreads due to liquidity providers exiting positions at zero profit (“scratch sales”) and also due to trading directly between non-liquidity traders (Smith and Whaley, 1994; Locke and Venkatesh, 1997).

A unique data set from the London International Financial Futures Exchange (LIFFE) that includes a complete record of bid and ask prices for two commodity futures markets is used in this chapter, in addition to the commonly available transaction price data. Use of the more complete LIFFE data thus facilitates an evaluation of the accuracy of spread estimates that might be computed when bid and ask data are not reported, as is the case in the large U.S. commodity futures markets. Accurate estimates of the nominal spread in these markets would give traders (and others) an idea of the “worst-case” transaction costs that they are likely to incur. In order to obtain a better descriptive evaluation of each estimator’s performance we test, for the first known time, for differences in the biases and variances of the spread estimators employing a procedure developed by Ashley et al. (1980). Indeed, this procedure allows us accurately isolate the strengths and weaknesses of each spread estimator. We also employ forecast encompassing techniques (Granger and Newbold (1973)), which reveal that there may be gains from combining estimates.

The second issue that we investigate here is the effect on spread magnitudes of moving from open outcry to electronic trading, which has been an issue of substantial controversy in recent years. It has been argued that electronic trading should be more efficient than other forms of trading, and many futures exchanges around the world have moved in this direction, either partially or fully. The advisability of the remaining open outcry futures markets moving to electronic trading remains the topic of intense debate, however, as some argue that the anonymity of such a system could result in increased rather decreased transaction costs (Massimb and Phelps (1994)). Given this interest, it is

not surprising that several studies have investigated the relative magnitudes of spreads in electronic and open outcry settings. Examples include Frino, McInish and Toner (1998), Wang (1999), and Tse and Zabolina (2001). These previous studies have investigated this issue with regard to *financial* futures markets, however, and there is no known study to date that has compared bid-ask spreads before and after a move to electronic trading in a *commodity* futures market. Commodity futures markets tend to have much lower trading volumes than financial futures markets, and have a relatively higher proportion of information traders (Foster and Viswanathan (1994)). Thus the automation of trading may have a different impact on spreads in a commodity futures markets than that in a financial futures market. A further contribution of this study is to evaluate the impact on nominal spread magnitudes of moving from open outcry to electronic trading in two LIFFE commodity futures markets, after controlling for spread determinants. The findings of this research have important implications for market participants and other exchanges that may be contemplating automating trading in their commodity futures markets.

The remainder of this chapter is organized as follows. In the following section we will assess the effectiveness of various spread estimators in estimating nominal spreads in commodity futures markets. After discussing the spread estimators and methodology that will be used in the evaluation, relevant data issues will be addressed and results will be presented. The following section evaluates the impact of the move from open outcry to automated trading on nominal spreads in commodity futures

markets, following a similar progression. Finally, we will offer some concluding remarks.

SPREAD ESTIMATOR PERFORMANCE

Bid and ask prices are usually not reported in open outcry futures markets and thus various estimators have been developed that estimate bid-ask spreads using commonly available transaction data. Naturally then, there has been an interest in assessing the performance of these estimators, but direct evaluation is made difficult by the fact that spreads are not observed (the very reason that made estimation of the spread necessary). Since direct evaluation has been difficult, researchers have argued the relative merits of spread estimators on theoretical grounds (e.g. Chu, Ding and Pyun, 1996), have compared estimates to expected patterns of spread behavior (Thompson and Waller, 1988), and have used simulations to evaluate estimator performance (Smith and Whaley 1994). To date, there have only been two direct evaluations of spread estimator performance. Locke and Venkatesh (1997) using clearinghouse records of scalper profits to directly evaluate estimator performance in estimating effective spreads in futures markets at the Chicago Mercantile Exchange (CME), finding that spreads estimators did a very poor job estimating effective spreads. Performances of spread estimators in the Financial Times Stock Exchange (FTSE) stock index futures market were directly evaluated by ap Gwilym and Thomas (in press), who found that estimators produced downwardly biased estimates of effective and nominal (quoted in their terminology) spreads.

The changes in transaction prices that are used to calculate spread estimates may be the result of "noise" trading, or the result of new information arriving in the marketplace. Different spread estimators employ various strategies to filter out the "true" price changes - those resulting from information arrival. It would seem reasonable therefore to believe that the relative proportions of these two types of trading in a market will have an impact on the accuracy of spread estimates. In commodity futures markets, there is a higher proportion of information traders than there tends to be in financial markets (Foster and Viswanathan, 1994). It is thus possible that the performance of a spread estimator in a financial futures market may not be indicative of that estimator's performance in a commodity futures market. It is for precisely this reason that we assess the effectiveness of various spread estimators in estimating *nominal* spreads in *commodity* futures markets. Accurate estimates of the nominal spreads in markets that do not report bid and ask data would be useful not only to market microstructure researchers, regulators, and exchange officials, but would give traders an idea of the "worst-case" transaction costs that they are likely to incur. Indeed, the bid-ask spread has an important impact on the profitability of trading activities, and failure to take the spread into consideration can lead to false conclusions in this regard (Bae et al., 1998; Shyy et al., 1996).

We now discuss the spread estimators and methodology used in this portion of the chapter. Spread estimators that have been developed in the literature have either utilized the covariance of successive price changes or have employed averages of absolute price changes. The former type of estimator, originally applied in equity

research, was first developed by Roll (1984). Roll made four assumptions, given which he developed a joint price distribution of price changes in a market that included market makers. First, he assumed an informationally efficient market. Second, he assumed that observed price changes had a stationary probability distribution. Third, he assumed that all customers made use of the market maker, who maintained a constant spread, s . Fourth, he assumed successive transactions would be sales or purchases with equal probability. Given these assumptions, he then deduces that any non-zero price changes that are not the result of the arrival of new information will be movements between the bid and ask prices, and any price change of zero is the result of two successive transactions at either the bid or the ask. This implied a joint probability distribution for successive price changes. He then calculated variances of price movements and the covariance of successive price movements (as functions of s), and proved that this calculated covariance conditional on no new information arriving was equal to the unconditional covariance of successive price changes. Solving the covariance equation for s resulted in Roll's estimator of the effective spread

$$RM = 2\sqrt{-\text{cov}(\Delta p_t, \Delta p_{t-1})} . \quad 2.1$$

Even though this estimator is intended to estimate effective spreads in equity markets, it is calculated and compared to observed nominal spreads in commodity futures markets in this study for purposes of comparison. This estimator has not typically been applied to futures transaction data because Roll's fourth assumption is often inappropriate for such data.

Chu, Ding, and Pyun (1996) suggested an estimator of the effective spread that relaxed Roll's fourth assumption that any given transaction has equal probability of taking place at the bid or the ask. They developed an estimator that incorporates the probability (δ) that an observed transaction takes place at the same price (bid or ask) as the previous transaction, and the probability (α) that an observed transaction takes place at the same price as the next transaction. These probabilities are estimated by applying a test, suggested by Lee and Ready (1991), that attempts to identify the price at which each transaction occurred. The reader is referred to Chu, Ding, and Pyun for the theoretical development of their estimator, as it is too lengthy to reproduce here. The resulting estimator is

$$CDP = \sqrt{\frac{-\text{cov}(\Delta p_t, \Delta p_{t-1})}{(1-\delta)(1-\alpha)}}. \quad 2.2$$

The estimators described thus far were designed with the intention of estimating effective spreads. Thompson and Waller (1988), however, proposed the following nominal spread estimator for futures markets:

$$TWM = \frac{1}{T} \sum_{t=1}^T |\Delta p_t|, \quad 2.3$$

where Δp_t , $t = 1, \dots, T$ is the series of non-zero price changes. They described this as being a function of the average bid-ask spread, and the magnitude and frequency of real price changes. Their estimator presumes that the average bid-ask spread component will be the primary determining factor, and no attempt is made to filter out real price changes. This estimator was applied in Thompson and Waller (1988) to study the

determinants of liquidity costs in feed grain futures markets, and was used to compare liquidity costs between two similar markets in Thompson, Eales, and Seibold (1988). Ma, Peterson, and Sears (1992) used the *TWM* to study intra-day patterns in spreads and the determinants of spreads for various Chicago Board of Trade (CBOT) contracts.

The estimator used by the CFTC to estimate the nominal bid-ask spread in futures markets was described in Wang, Yau, and Baptiste (1997). Like *TWM*, this estimator also takes an average of absolute non-zero price changes, but attempts to remove the effect of real price changes by omitting any price change that follows another price change of the same sign. That is to say, the CFTC estimator is the average, absolute, *opposite direction*, non-zero price change. This requirement that some data be omitted means that a greater quantity of data may be required to calculate a spread estimate. In thinly traded markets, “bounces” between the bid and ask prices may be fairly infrequent while real price changes may be more numerous.²

Smith and Whaley (1994) adopted a different strategy to account for the effects of true price changes when estimating futures market spreads. They made two assumptions. First, they assumed that the spread is constant over the time frame for which it is being estimated. Second, they assumed that the expected value of *true* price changes is zero. They did not assume, however, that the variance of true price changes is zero, an assumption in *TWM*. Then, taken as given that the observed price series does

² Another estimator, proposed by Bhattacharya (1983), is the average of an even smaller subset of absolute price changes. Because the markets considered here have fairly low volumes except in the contracts nearest delivery, this estimator would have frequently not produced an estimate. Those interested in estimating nominal spreads for higher volume commodities or contracts may wish to consider this estimator.

not include repeated observations of the same price, they derived the first and second population moments of the *observed* price changes. These are functions of both the spread and the variance of true price changes. These population moments were then set equal to the sample moments of the observed price changes, and these two equations were solved for the two variables. Hence Smith and Whaley arrived at an estimator for the effective spread that explicitly accounts for the effects of true price changes.

Given a set of available estimators and observations of nominal spreads, we must determine the statistical methodologies to be used in assessing estimator performance. One simple method might be to test for equality of the means of squared errors, or some other measure of economic loss, for each pair of two estimators using a simple *t*-test procedure. However, in order to get a better descriptive evaluation of the performance of each estimator, here we test for differences in the biases, variances of the estimators using a procedure developed by Ashley et. al (1980).

Specifically, from the definition of mean squared error, it is simple to show that for two forecasts with errors e_1 and e_2 that:

$$MSE(e_1) - MSE(e_2) = [s^2(e_1) - s^2(e_2)] + [m(e_1)^2 - m(e_2)^2], \quad 2.4$$

where MSE is the sample mean square error, s^2 is the sample variance, and m is the sample mean error. Defining:

$$\Delta_n = e_{1n} - e_{2n} \text{ and } \Sigma_n = e_{1n} + e_{2n}, \quad 2.5$$

then equation 2.4 can be rewritten as:

$$MSE(e_1) - MSE(e_2) = [\text{cov}(\Delta, \Sigma)] + [m(e_1)^2 - m(e_2)^2]. \quad 2.6$$

The null hypothesis that there is no difference in the mean squared error of two estimators is then equivalent to the null hypothesis that both terms on the right hand side of equation 2.6 are zero. This can be tested by regressing:

$$\Delta_i = \beta_0 + \beta_1 [\Sigma_i - m(\Sigma_i)] + u_i . \quad 2.7$$

This results in least squares estimates:

$$\hat{\beta}_0 = m(e_1) - m(e_2) , \quad 2.8$$

and

$$\hat{\beta}_1 = [s^2(e_1) - s^2(e_2)] / s^2(\Sigma) . \quad 2.9$$

Testing that both terms on the right hand side of equation 2.6 are zero is equivalent to testing $\beta_0 = \beta_1 = 0$. If either of the two least squares coefficient estimates is significantly negative, the null hypothesis that the *MSE*'s are equal automatically cannot be rejected. If one coefficient estimate is negative but not significantly so, a one-tailed t-test on the other estimate can be used. If both estimates are positive, then an *F*-test that both coefficients are zero can be performed, but a significance level equal to one quarter of the usual level must be used (Brandt and Bessler 1983).

In addition to allowing a test of the null hypothesis that two *MSE*'s are equal, estimating equation 2.7 also facilitates testing whether or not the biases and variances of two estimators are equal. From equation 2.8, it is obvious that an estimate of β_0 that is significantly different from zero implies that the two biases are different. Similarly, an estimate of β_1 significantly different from zero implies that the two variances are

different. Equation 2.7 is estimated for each combination of two estimators for each commodity in this study to test for equality of their biases and variances.

In addition to testing the biases and error variances of estimators against one another, we also test whether or not any of the estimators are redundant (i.e. contain no unique information). This is essentially the idea behind encompassing, which is closely related to conditional misspecification analysis and composite forecasting. In particular, Granger and Newbold (1973) suggested the use of a composite estimator.

$$E_{cn} = (1 - \lambda)E_{1n} + \lambda E_{2n} , \quad 2.10$$

where E_{1n} and E_{2n} are two component estimators and $\lambda \in [0,1]$ is a parameter to be estimated. The error of this composite estimator is equal to the error of the first component estimator plus λ multiplied by the difference of the errors of the two components. Thus the equation:

$$e_{1n} = \lambda(e_{1n} - e_{2n}) + u_n , \quad 2.11$$

can be estimated to determine if estimator 2 contains information not present in estimator 1 (Harvey et al. 1998). If $\lambda = 0$ cannot be rejected, then estimator 2 does not contain any additional useful information, and estimator 1 is said to “encompass” estimator 2. Therefore, in this study, equation 2.11 is estimated for each permutation of two estimators for each commodity, to determine if any of the estimators are redundant. As suggested by Harvey et al. (1998), White’s heteroskedasticity-consistent variance of the estimate of λ is used, as the error series e_{in} exhibits skewness and kurtosis that strongly suggest a non-normal distribution for each estimator i .

We now describe the data used to evaluate spread estimator performance. All bid, ask and transaction prices for cocoa and coffee futures contracts are provided by LIFFE on the “LIFFEstyle 2000” data CD. This stands in contrast to the major U.S. futures exchanges, where transactions at the price of the previous transaction are not reported, and bids and asks are only reported when little nominal trading is occurring (Locke and Venkatesh 1997). Before July 3rd 2000, these data were time-stamped only to the nearest minute, making the construction of nominal spreads by matching contemporaneous bidding and asking prices an imprecise exercise. As such, these data are not used in the present study. However, from July 3rd 2000 through November 27th 2000, the bid, ask and trade prices generated during open outcry trading were time-stamped to the nearest second. The data generated during this period of time thus facilitate the accurate matching of contemporaneous bidding and asking prices, and the differences between these prices constitute nominal spread observations.

The LIFFE cocoa contract calls for delivery of 10 tonnes (metric tons) of cocoa, with a minimum price fluctuation of one pound sterling per tonne. Delivery months are March, May, July, September, and December. The daily volume of trading in the nearby futures averages about 1446 contracts per day over the time period from July 3rd 2000 through November 27th 2000. LIFFE coffee futures contracts call for delivery of 5 tonnes of robusta coffee. The minimum price fluctuation is one U.S. dollar per tonne, and delivery months are January, March, May, July, September, and November. Daily trading volume in the nearby futures is roughly 1985 contracts. Examples of the data

reported for November 2000 coffee futures on 27 September 2000 are provided in Table 2.1.³

As previously noted, bid and ask prices are not necessarily called out simultaneously by a single trader. Observations of the bid-ask spread for each market are thus constructed by matching a bid or ask price with a price of the opposite type that occurred within a chosen time interval. Bid and ask prices called out in open-outcry futures trading are only required to be honored if they are immediately accepted by another trader, although it has been noted that in practice traders (especially scalpers) let their bids and offers “live” (Silber 1984).⁴ Thus the choice of the time interval used to construct spread observations presents a tradeoff. A relatively restrictive time criterion naturally result in fewer spread observations, but one can be more assured that these observations represent a valid nominal spread. A less restrictive criterion results in more observations, but some of these observations may be too far apart in time to have constituted a valid nominal spread.

A second, related criterion must be considered. The resulting spread observations are then used to calculate daily average spreads. In order to ensure that a given daily average is in fact representative of the spreads that prevailed on that day, one must insist on some minimum number of spreads to use in calculating that average.

³ All data are subjected to a screening algorithm and obviously erroneous observations are removed.

⁴ This stands in contrast to electronic data whereby any bids or asks that are reported by the exchange are standing limit orders and will exist until the trader actively withdraws the bid or ask. As such, the bid-ask data series from an electronic trading environment looks very different than that from an open outcry environment.

In this research, the highest quality of observations (shorter time interval for spreads, more spreads per day when constructing a daily average) is used that still allows an acceptable quantity of observations for reliable statistical analysis. Specifically, a 10-second maximum time interval is used for constructing a spread, and a minimum of 20 spreads are used for calculating a daily average.⁵ Varying these criteria somewhat does not result in significant changes to the qualitative results reported below. Applying the 10-second criterion to the data in table 2.1, bid-ask spreads of \$1 per tonne are observed at 10:04 a.m. and 10:18 a.m.

Since we are comparing estimates of the daily average spread to observations of the daily average spread, it is advisable to be sure that that daily average is generally representative of spreads observed throughout the day. In high-volume financial futures markets, there are well-documented intra-day patterns in bid-ask spreads (e.g., Tse, 1999). It is therefore possible that calculating an average spread over the length of a day in this application might “mask” a consistent pattern of significant intra-day deviations of spreads away from the overall daily average. In order to check for such a phenomenon, the trading day was divided into six roughly equal length time frames for

⁵ Prices must be successive. For example, suppose a bid occurs at 10:00:00, and another, different bid occurs at 10:00:03. Then, an ask is observed at 10:00:07. This ask would not be mated with the first bid, even though they both occurred within 10 seconds of one another. In an earlier version of the research the same analysis was conducted on the open outcry trade data provided by LIFFE from 1996 to July 2000 (before the reporting system changed). As mentioned previously, this data series meant that many of the bids and asks reported within the same minute did not represent a valid spread (e.g., non positive spreads) and so did not represent the true course of events within that minute. Results from this analysis, that excluded these non-positive spreads were not entirely dissimilar to the results presented in this chapter and are excluded to conserve space.

each commodity, and average nominal spreads were calculated for these intra-day time periods for each day. The deviation of each intra-day average spread from the overall daily average was then calculated for all days in the sample. The null hypothesis that the mean deviation for an intra-day period was significantly different from zero was tested, and results of these tests are reported in table 2.2. We find that on average, the first period average spread is above the daily average (as reflected by the negative deviations reported in table 2.2), and generally the subsequent periods' average spreads are below the daily average. This suggests a weak “reverse-J” pattern of spreads similar to that found in ap Gwilym and Thomas (in press). However, none of the intra-day spread deviations were found to be significantly different from zero, implying that there is no consistent pattern of significant intra-day deviation of commodity futures spreads away from the daily average spreads over the sample period. We can thus feel comfortable in following the significant body of research that has employed estimates of the daily average spread, and do not apply the estimators to shorter time frames.⁶

The daily spread averages for a contract in our data sample generally follow a “U-shaped” pattern in which they are higher when the delivery date is distant, decrease as time passes, and eventually increase as the delivery date approaches. As an example, spreads for the November 2000 coffee contract are plotted over time in figure 2.1.

The transaction observations provided by LIFFE include consecutive transactions at equal prices. From this data, a “raw” series of price changes is constructed, which is

⁶ Indeed, spread estimators have been used to estimate spreads over even longer time periods. For instance, Laux and Senchack (1992) estimated monthly average spreads in financial futures markets and used these estimates to carry out their analysis.

then used in the calculation of RM . It should be noted that this type of transaction price series is not reported by the major U.S. exchanges, and so the RM estimator could not be applied to U.S. data in the way that it is applied here. A series consisting of strictly non-zero price changes is constructed, which is then used to calculate CDP , TWM , and SW . This second price change series is thus like that which would be reported by a U.S. futures exchange. Lastly, a series of only opposite-direction price changes is assembled for use in calculating $CFTC$. This last price change series typically contains about half as many price changes as the strictly non-zero price change series, which in turn usually contains about half as many price changes as the completely unrestricted price change series.

We come now to the evaluation of the spread estimators' performances. The daily average bid-ask spread is estimated for each day of each delivery over the time period from 3 July 2000 through 24 November 2000. The serial covariance-type estimates, RM and CDP , frequently cannot be calculated however due to price changes that exhibit positive serial covariance. This occurs relatively more often for cocoa (about 44% of observations) than for coffee (about 20% of observations). Within each commodity the problem occurs more often for CDP , the serial covariance estimator using only price-changing observations. Other researchers have noted this problem with serial covariance estimators and have offered various explanations. For instance, Chu, Ding, and Pyun suggested that positive serial covariance in price changes could be due to sequential information arrival. Roll suggested that market inefficiencies over short

time frames could be to blame. Observations where *RM* and *CDP* encounter the problems described above are omitted from the analysis.

Correlations between observed and estimated daily average spreads for each market are given in table 2.3. All of the estimates are more highly correlated with the observed spreads for coffee than for cocoa, with the exception of *RM*. The correlations between the serial covariance estimates and the observed average spreads are positive, but not especially high, ranging between 0.12 and 0.32. Correlations between the remaining estimates and average spreads are more impressive, falling in the 0.47 to 0.85 range. In this respect, *TWM*, *SW*, and *CFTC* appear to do a much better job than *RM* and *CDP*. *TWM*, *SW*, and *CFTC* are highly correlated with one another, as are *RM* and *CDP*. Thus estimators of the same type (serial covariance-type estimators or absolute price change-type estimators) seem to be highly correlated with one another, and noticeably less correlated with estimators of the other type. Interestingly, even though *SW* is designed to estimate effective spreads, it is more highly correlated with the nominal spread estimators (*TWM* and *CFTC*) than with the other effective spread estimators (*RM* and *CDP*).

Performance of the estimators using various measures for all observations are given for each commodity individually in table 2.4. The performance of the estimators *relative* to one another is similar within each commodity. The absolute price change-type estimators seem to perform much better than the serial covariance type estimators by each of the performance measures. Among the absolute price change estimators, relative performance is very similar for cocoa. However *SW* performs somewhat worse

than *TWM* and *CFTC* when estimating coffee spreads. Thus the relative performance *SW* estimator may be somewhat inconsistent across commodities.

Comparing the *absolute* performance of the estimators across commodities using the mean absolute percent error measure, the absolute price change estimators seem to perform somewhat worse when estimating coffee spreads than when estimating cocoa spreads. We also note that all mean errors are negative for all estimators for both commodities, suggesting that the estimators produce downwardly biased estimates of nominal spreads in commodity futures markets. This is consistent with the findings of ap Gwilym and Thomas for financial futures.

The results from the estimation of equation 2.7 for each combination of commodities are presented in table 2.5. In almost all cases, the null hypotheses that $\beta_0 = 0$ is rejected at the 5% level of significance, meaning that for the most part the differences in the biases (mean errors) reported in table 2.4 are significant. The sole exception is that the difference in the biases of *TWM* and *CFTC* for cocoa is not significantly different. In most cases the null hypothesis $\beta_l = 0$ also cannot be rejected, with the interesting exceptions being that the error variances of *TWM* and *SW* are not significantly different for cocoa, and the error variances of *CFTC* and *TWM* are not significantly different for coffee.

It should be noted at this point that all results reported thus far are based on all data for all contracts. The u-shaped pattern in figure 2.1 suggests that conditions over the life of a contract vary, and thus performance of spread estimators may thus vary by time to delivery. However, only the aggregate results are only presented as separating

the data into nearby and distant groups revealed only a single interesting difference in performance. This difference is that for cocoa, the bias of the *CFTC* estimator improved to be significantly better than the *TWM* estimator, and the variance of the *CFTC* estimator improved to be not significantly different from the *SW* and *TWM* estimators. Thus the performance of the *CFTC* estimator may be somewhat better when estimating spreads for a contract nearby delivery.

Analyzing the signs of the coefficient estimates in table 2.5, the serial covariance estimators have larger biases than the absolute price change estimators (significantly positive β_0 estimates), but lower error variances (significantly negative β_1 estimates). This naturally leads one to question which class of estimators generally has lower means of squared errors. As discussed earlier, in some cases an F-test can be used to test the null hypothesis that both β_0 and β_1 from equation 2.7 are zero for a pair of estimators, implying that the mean squared errors of the two estimators are not significantly different. However if one of the two coefficient estimates is significantly negative, this null hypothesis automatically cannot be rejected. This is the case for most of the possible pairs of estimators in this study, and thus the Ashley methodology is largely powerless for finding differences in the mean squared errors here. Although the statistical methodology available cannot prove that the means of the squared errors of the serial covariance estimators are greater than those of the absolute price change estimators, the relative magnitudes reported in table 2.4 strongly suggest that this is the case. Still, these results suggest that those interested in minimizing error variance (at the

expense of significantly higher error bias) may wish to consider using the serial covariance estimators.

The other methodology we employ to evaluate the estimator performance is the forecast encompassing testing procedure described previously. Probability values for the test that $\lambda = 0$ (from equation 2.11) for each permutation of two estimators are presented in table 2.6. In most cases, the null hypothesis that one estimator encompasses another is rejected. In only one case is this hypothesis not rejected across both commodities: we cannot reject that *CDP* encompasses *RM*. Since encompassing is generally rejected, it is quite possible that a composite estimator could provide superior estimates of nominal bid-ask spreads. In particular, one might speculate that combining a serial covariance estimator and an absolute price change estimator might prove fruitful, as the former will have a lower error variance, while the latter will be less biased.

SPREADS IN ELECTRONIC AND OPEN OUTCRY COMMODITY FUTURES MARKETS

It seems therefore that spread estimators may be useful for traders not able to consistently observe bidding and asking prices, as on U.S. exchanges. Indeed, as mentioned previously, spread estimators may shed some light on likely transaction costs. However, of late many trading environments have moved toward automated trading, suggesting that the costs of trading may indeed alter. Whether or not moving to an electronic platform affects bid-ask spreads in commodity futures markets is a question to which we now turn.

Arguments on the relative merits of open outcry and automated trading systems have focused on two issues. First, researchers have noted that a market maker faces an adverse selection problem (Copeland & Galai (1983), Glosten & Milgrom (1985)). Specifically, if a market maker must make a commitment to buying and selling prices that will be available to all traders, she exposes herself to counterparties with superior information. Features of the open-outcry system mitigate the severity of the adverse selection problem to some extent, however. In the open-outcry environment, traders are face-to-face with their counterparties, and can thus infer from their identity and disposition the likely nature of the information that they possess. Furthermore, if they perceive that private information might be entering the market, traders can very rapidly adjust their offers to buy and/or sell. In an anonymous limit order book system, however, the market maker is deprived of these advantages. As noted by Copeland & Galai (1983), a limit order can be likened to a short option position with a time to maturity equal to the time required to withdraw the order. In an anonymous limit order book system the market maker is forced to make this option available to all traders, and will not be able to quickly discern when the well-informed are entering the market. Market makers will thus require compensation, in the form of wider bid-ask spreads, for this increased risk that they will be at an informational disadvantage on any given trade. It is thus widely believed that a more pronounced adverse selection problem will tend to increase transaction costs in anonymous electronic trading, relative to open outcry trading.

The adverse selection problem may be more acute in some markets than in others, however. The model of Subrahmanyam (1991) suggests that the information costs paid in a market for a basket of assets (e.g. a stock index futures market) are lower than those paid in a market for an individual asset. Also, the values of some assets are determined largely by information that is naturally public in nature. For example, the values of debt instruments are likely to be a function primarily of the state of the macroeconomy, which is relatively easily observed in countries that report a comprehensive set of national accounts. Prices in other markets, however, are likely to be determined by information that is held by a relatively small number of agents (e.g. an agricultural commodity market). In these types of markets, Foster & Viswanathan (1994) suggest that the well-informed traders will indeed capitalize on their advantageous position. Thus the impact on transaction costs of moving from open outcry to automated trading is likely to depend on the specific nature of the market, and the results found in financial markets may not apply to commodity futures markets.

The order processing component of transaction costs are the other issue on which the relative merits of open outcry and automated systems have been compared. It has been suggested that automated trading should offer significant operational efficiencies relative to open outcry trading, thereby potentially reducing transaction costs (Massimb and Phelps (1994), Pirrong (1996)). Specifically, fewer people need be employed in an electronic system, and electronic trading should result in fewer costly mistakes (out-trades) than open-outcry trading. A significant fixed cost is likely to be associated with automating trading, however, and there may be much less potential for gains in

efficiency in a fairly low volume futures market. On numerous levels therefore we have reason to believe that results regarding the impact on spread magnitudes of automating trading found in financial futures markets may not apply to commodity futures markets.

Despite the possible differences in impact it is worthwhile providing a very brief (and by no means comprehensive) summary of some results found in the financial markets. However, the summary is by no means comprehensive. Frino, McInish, and Toner (1998) (among others) examined simultaneous electronic and open outcry trading in German Bund futures, finding wider spreads on the automated exchange. They also found that during electronic trading, there was a larger marginal effect of an increase in volatility on spread magnitudes. Wang (1999) analyzed the differences between daytime open outcry and evening electronic trading in financial futures at an exchange, finding results similar to Frino, McInish, and Toner. Tse and Zobotina (2001) looked at trading in FTSE stock index futures before and after trading was automated at LIFFE, finding that spreads were narrower after the trading was automated. Thus the evidence regarding the relative magnitudes of bid-ask spreads in electronic and open outcry financial futures markets is mixed.

Here, we will compare the magnitudes of bid-ask spreads before and after a move from open outcry to automated trading in the same two commodity futures markets evaluated earlier (cocoa and coffee). Compared to the financial futures markets examined in the studies cited above, these markets have significantly lower trading

volumes.⁷ These lower volumes call into question the potential for increasing operational efficiency by automating trading. Also, for reasons discussed earlier, the impact of moving to electronic trading on the adverse selection component of transaction costs is likely to be different for these markets than it is for the financial markets studied previously. The relative proportions of well-informed traders are different in commodity futures markets than in financial markets, and the information that determines prices in these markets is inherently less public in nature. We posit the following hypotheses:

Hypothesis 1: In anonymous electronic trading, bid-ask spreads have a greater tendency to widen in response to increases in volatility (relative to open outcry trading).

Hypothesis 2: Market makers will face a significant adverse selection problem in an anonymous electronic commodity futures market, and net transaction costs, as measured by bid-ask spreads, will be higher than those observed in the open outcry system.

We first describe the methodology used for comparing electronic and open outcry spreads. We will use the methodology employed by Frino, McInish, and Toner to compare spreads on a security traded at two different exchanges while controlling for factors known to affect spreads. Rather than comparing spreads at two different exchanges, however, we will be comparing spreads before and after a switch from open-outcry to electronic trading, as in Tse and Zabortina. The empirical model is as follows:

⁷ For example, between July 2000 and June 2001, volume on the FTSE 100 futures contract at LIFFE was 9,033,641, whereas volumes on the cocoa and coffee markets were 1,408,945 and 1,271,816 respectively.

$$\begin{aligned}
Spread_t = & \beta_0 + \beta_1 D_e + \beta_2 \sqrt{volume_t} + \beta_3 \sqrt{var_t(price)} + \beta_4 D_e \sqrt{volume_t} \\
& + \beta_5 D_e \sqrt{var_t(price)} + \beta_6 \sqrt{price_t} + \varepsilon_t.
\end{aligned}
\tag{2.12}$$

$Spread_t$ is the average nominal bid-ask spread during period t , D_e is a dummy variable that is zero for an open-outcry observation and one for an electronic observation, $volume_t$ is the total volume of futures traded, $var_t(price)$ is the variance of spread midpoints during period t , and $price_t$ is the average spread midpoint during period t . Consistent with McNish and Wood (1992) and Frino, McNish, and Toner, square roots of the determinants of the spread are used to prevent outlying observations from exerting undue influence on the regression results. Theory suggests that we should expect a negative relationship between spread magnitude and volume of trade, and a positive relationship between spread magnitudes and price variability (Copeland and Galai (1983)). These results have also been observed in empirical studies (examples include McNish and Wood (1992) for stocks, and Ding (1999) for futures). The relationship between quoted spreads and the level of the price of the commodity is expected to be positive for two reasons. First, the volatility of prices of commodities tends to increase as the prices themselves increase. Thus it is possible that the price coefficient in the model might “pick up” some of the positive effect of price variance on spreads. A similar argument was used by Stoll (1978).⁸ Second, and perhaps more importantly, it is

⁸ Stoll was dealing with stocks rather than commodities. He argued that the risk associated with a stock *decreased* as the price of a stock increased, and therefore he expected to find a *negative* relationship between spreads and the price level of the underlying security. Here we use a similar argument, but expect to find an effect opposite to that found by Stoll due to the differences in the instruments under consideration.

expected that *percentage* spreads should be somewhat steady. In the words of Demsetz (1968), a positive relationship is expected between nominal spreads and the price level so as to “equalize the cost of transacting per dollar exchanged.”

We now describe the data used in comparing electronic and open outcry Spreads. Nominal spread observations must be constructed differently when using the electronic trading record rather than the open outcry trading record, taking into consideration the different trading mechanisms. In the electronic system at LIFFE, bid and ask price observations are the result of standing limit orders, and need not be acted upon immediately as in open-outcry trading. This essentially means that there is now a spread observation at every point in time during the trading day. For each trading day from 27 November 2000 through 11 May 2001, a time series of observations of the prevailing spread for each second was constructed for the nearby futures. A daily average spread was then calculated by averaging over the observations for each second.

Before controlling for the determinants of spreads, we find similar average daily spreads for nearby coffee futures in the electronic and open-outcry periods, at \$1.97 and \$2.07 per tonne, respectively. We observe a noticeable increase in daily average spreads for cocoa, however. Over the open-outcry period, the average spread for nearby futures is £1.56 per tonne. In the electronic period, the average spread is a noticeably higher £3.31 per tonne, and there is a much greater variability relative to the open-outcry period. Cocoa prices experienced a significant increase (which is usually accompanied by an increase in volatility) shortly after the move to automated trading, however, and it is therefore important to control for such factors before drawing any conclusions.

We now present the results of comparing electronic and open outcry spreads. The model in equation 2.12 was estimated for each of the two commodity futures markets, and robust standard errors for the parameter estimates were estimated using the Newey and West (1987) procedure. Results are presented in table 2.7. For cocoa, we find that the volume and volatility coefficients not significant. This is somewhat surprising, although these results may be due to the fact that there was little variation in these variables (and indeed the dependent variable itself) during the open-outcry period. While not significant, the coefficient on the price standard deviation term has the expected sign. The price level term has a positive coefficient, as expected, and is highly significant. Also, we find a significantly negative constant term. Although this may seem counter-intuitive given that the dependent variable is always positive, the results must be taken as a whole. The square root the price variable has a mean of 26.3, with a standard deviation of 2.1. Thus the highly significant coefficient on this variable of 0.329 implies that this term is consistently adding about £8 to the predicted spread. This suggests that the negative constant is no cause for concern.

The coefficient on the electronic dummy variable is positive and significant at the 10% level. We therefore find that, after controlling for other explanatory factors, the switch to electronic trading has widened observed spreads in the cocoa futures market by about £0.64 per tonne. The coefficient for the volume interaction term is negative and significant. This indicates that cocoa spreads have become sensitive to volume since the move to electronic trading. Increases in the volume of trade cause decreases in the spread, whereas no such effect was observed during the open-outcry period. We also

find a significantly positive volatility interaction term, suggesting that cocoa spreads have also become sensitive to volatility following the move to automated trading.

Turning to the coffee results, we find coefficients of the expected signs for the volume and volatility terms, with the volatility term being significant. As in the cocoa model, the coefficient on the price level is positive and significant. Also as in the cocoa regression, we find a positive and significant electronic dummy term, a positive and significant volatility interaction term, but no significant volume interaction term. Thus, as in the cocoa market, we find that spreads in the coffee market have become more sensitive to the level of price variability than they were during open outcry trading, and have generally widened after controlling for spread determinants.

In both markets we find the result that transaction costs, as measured by the magnitudes of bid-ask spreads, have a greater tendency to increase as prices become more volatile, supporting our Hypothesis 1. This observation is consistent with the suggestion that market makers in the anonymous automated market cannot distinguish between noise trading and information trading. They thus have an increased tendency to widen spreads during high-volatility periods as compensation for the risk that they may be at an informational disadvantage. This result is consistent with results from financial futures research (Frino, McInish, and Toner (1998) and Wang (1999)).

The finding that spreads have widened in the cocoa and coffee futures markets suggests that the net effect of automating trading has been to increase transaction costs. We thus find support for Hypothesis 2. Specifically, these results suggest that lower order processing costs are outweighed by increases in transaction costs due to a more

severe adverse selection problem. This suggests that one of the expected benefits of electronic trading, reduced transaction costs as manifested by narrower bid-ask spreads, may not materialize, depending on the nature of the market in question. Commodity futures markets in particular, with their lower volumes and higher proportions of information traders, may not realize lower transaction costs by automating trading.

Given that we have found that the size of the spread has changed with the change in environment a critical question is that of the economic significance of the differences in spreads observed since the move to electronic trading. Indeed, from both a market participant and exchange point of view having an understanding of the monetary implications of executing a trade in the electronic environment is paramount. Therefore, similar to an analysis carried out in Venkataraman (2001), we use our empirical models of coffee and cocoa spreads to calculate the potential increases in transaction costs that have been realized since trading was automated. Specifically, we calculate the estimated impact on the spread due to the automation of trading at time t as

$$Change_t = \hat{\beta}_1 + \hat{\beta}_4 \sqrt{volume_t} + \hat{\beta}_5 \sqrt{var_t(price)} \quad 2.13$$

for each commodity, where the coefficient are from the appropriate estimate of equation 2.12. Note that this represents the estimated average change in the spread per ton on a particular day. We then multiply this number by the number of tons in the contract to arrive at an estimated change in the spread per contract. This value is then averaged over the entire electronic trading period in our sample, weighting each day's observation using that day's volume. We calculate these values as £6.46 for cocoa and \$3.94 for coffee. These numbers might be interpreted roughly as the average increase (due the

automation of trading) in transaction cost per contract that is being realized by a trader who completes a round-turn using market orders to both enter and exit the position. Care must be exercised in this interpretation, however, as these are nominal spreads rather than effective spreads.⁹ Nonetheless, these numbers give some sense of the economic impact of the move to automated trading in these commodity futures markets and illustrate that for the commodity markets studied here, the change in environment has increased transaction costs.

CONCLUSIONS

This study has investigated issues regarding nominal bid-ask spreads in relatively low-volume commodity futures markets. Several spread estimators were applied using open outcry transaction data from the LIFFE coffee and cocoa markets, and the resulting estimates were compared to observed nominal bid-ask spreads. The mean absolute price change estimators, *TWM*, *CFTC*, and *SW*, perform better at estimating daily average nominal spreads than the serial covariance estimators, *RM* and *CDP*, by the bias and mean square error criteria. The serial covariance estimators have lower error variances, however. Encompassing test results generally confirm that the estimators do not encompass one another, and there may be gains from combining estimates. These

⁹ This interpretation of the nominal spread is safe for the automated market, as a market order cannot be executed within the prevailing nominal spread. In the open outcry market, traders entering market orders may have enjoyed effective spreads that were lower than nominal spreads. This implies, however, the measures of the economic significance of the wider spreads that we calculate and interpret can be considered conservative.

results should be of interest to those who wish to estimate potential transaction costs in open outcry futures markets that report transaction price data, but not bid and ask data.

We find an increased tendency for spreads to widen as volatility increases, which is consistent with the argument that market makers face a worse adverse selection problem in anonymous electronic trading. Also, we find that net transaction costs, as measured by bid-ask spreads, have widened in the commodity futures markets studied here, even after controlling for spread determinants. This suggests that lower order processing costs in automated trading may be outweighed by increases in transaction costs due to a more severe adverse selection problem. It thus seems that some of the benefits that have been realized by automating trading in some financial futures markets may not be realized in commodity futures markets, which tend to have lower volumes and are inherently different in nature.

CHAPTER III

CAUSALITY IN FUTURES MARKETS

INTRODUCTION

It has been over seventy years since Keynes wrote his *Treatise on Money*, in which he proposed his theory of “normal backwardation” – the idea that hedgers use futures markets to transfer risk to speculators, causing futures prices to deviate from expected future cash prices so that the speculators might be compensated. Despite decades of empirical investigation, no consensus regarding the validity of Keynes’ conjecture has been reached. Two difficulties have prevented the conclusive confirmation or rejection of the theory. First, the expected future cash price is not observed, and therefore neither is any risk premium. Second, it is not feasible for researchers to seek the answer to this question by experimentation. Systematic manipulation of futures markets is not only impractical; in many cases it is illegal.

Indeed, researchers conducting empirical work in economics and finance generally must work with observational rather than experimental data, and frequently are not able to observe all relevant quantities. This makes the correct inference of causal relationships difficult at best and impossible by some accounts - many assume that the use of controlled experiments is the *only* means by which causal mechanisms can reliably be inferred. Careful empirical researchers in these fields have thus resigned themselves to being able to draw only rather weak conclusions. It is said that evidence consistent with a theory is found, rather than that a theory has been proven. The less

cautious investigator, upon finding that two observed quantities A and B covary, might imprudently conclude that A causes B , or vice versa. Consequently, empirical studies in economics and finance rarely unanimously support or reject available theoretical explanations. Such is certainly the case with research into futures markets, the subject of this dissertation.

This situation is not unique to economics and finance - researchers in numerous fields find themselves operating under such difficult circumstances. This situation has inspired a recent multidisciplinary effort to develop a body of theory concerning the inference of causal relationships using observational data. A subset of this literature further concerns itself with conducting this inference when the observational data are incomplete. Treatments of this subject can be found in Pearl (2000) and Spirtes, Glymour and Scheines (2000). This study uses these causality methods to investigate Keynes's theory, and other unresolved questions of more recent vintage regarding futures markets: we investigate the causes of the well-documented positive correlation between volume and volatility in futures markets, and assess the evidence regarding theories that predict that the activities of certain types of traders affect levels of price volatility. A correct understanding of the causal mechanisms that drive futures markets is obviously important for a variety of parties – hedgers, speculators, exchange officials, and regulators.

The following section extends this introduction by discussing in detail the issues that we investigate and the importance of each. We then describe the specific causal

inference procedure that we employ and the data that we use. Finally, we present the analysis and offer some concluding remarks.

ISSUES INVESTIGATED

The first issue that we investigate is the Keynes's (1930) theory of normal backwardation, and its extensions. Keynes believes that hedgers enter the futures markets primarily to reduce the risk associated with cash market positions. He further believes that hedgers are generally commodity producers, and are therefore long in the cash market and short in the futures market. This necessarily means that speculators must be long in the futures market, and he postulates that the current futures price must be below the expected future cash price in order to induce the speculators to bear the risk associated with those long positions. A consequence of the theory as stated by Keynes, and approved by Hicks (1939), is that a futures contract's prices are expected to display an upward trend on average. Telser (1958) searches for such trends in the cotton and wheat markets, and reports finding no evidence. Cootner (1960) extends the theory by pointing out that hedgers are not necessarily commodity producers, but may be commodity consumers as well. Thus the net position of hedgers as a whole might be either long or short. As such, he suggests that the current futures price might be either above or below the expected future cash price. The modified theory is sometimes referred to as "net hedging" or "hedging pressure." Cootner reports finding evidence consistent with this hypothesis, namely that speculators appear to be earning profits over his sample period. Cootner thus shifts the empirical focus from searching for trends in

futures prices to searching for speculative profits and/or hedging losses in futures markets.

Houthakker (1957) and Rockwell (1967) note however that speculative profits may be due to superior forecasting ability, rather than the collection of a risk premium. Rockwell thus recasts normal backwardation as “the return earned by a hypothetical speculator who follows a naïve strategy of being constantly long when hedgers are net short and constantly short when hedgers are net long.” This then implies that speculative profits / hedging losses are a necessary, but not sufficient condition for the net hedging theory to hold. The analysis then must focus on decomposing speculative profits into forecasting ability and naïve components. Both Rockwell and Chang (1985) conduct such analyses, and each finds evidence of speculative profits. Rockwell reports that speculative profits are due to forecasting ability, however Chang reports evidence of naïve profits as well. This approach suffers from the inherent difficulty of dividing speculators into able and naïve groups, when forecasting ability is unobserved. The reliability with which this task can be performed using aggregate data on trader positions is highly questionable. Further complicating matters, the available data regarding market commitments contain a proportion of traders whose status (either speculator or hedger) is unknown. In contrast to Rockwell and Chang, Hartzmark (1987) uses a very unique, highly disaggregated data set to find evidence that hedgers earn significant positive profits on average, precluding Rockwell’s naïve speculator from profiting. Certainly, the evidence from these related empirical approaches to the question is mixed.

Meanwhile, another thread of the literature has developed a somewhat different perspective on the question. The theory normal backwardation is presented in the context of a single asset, and the hypothesized risk premium should therefore due to the expected futures return and variability of that return. Dusak (1973) and Black (1976) argue that the question should be considered in a portfolio context. The capital asset pricing model (CAPM) states that any risk premium should be due to the relationship between an asset's returns and returns on total wealth. If a futures contract's price changes are correlated with returns on total wealth, then some portion of the risk of holding a futures contract is undiversifiable (a non-zero "beta" in CAPM parlance), and a risk premium should therefore be present (because there is a "systematic" risk). If, on the other hand, futures price changes are independent of returns on total wealth, then the risk of holding a futures contract should be fully diversifiable, and no risk premium should be present. Dusak finds that for the markets that she considers, futures price changes are independent of returns on a proxy for total wealth (the S&P 500 index), and concludes that no risk premiums are present. Carter, Rausser & Schmitz (1983) argue that Dusak's proxy for total wealth is inadequate, and that it should include commodity prices. Hirschleifer (1988) and Hirschleifer (1990) argue that the assumptions built into the standard CAPM might be inappropriate, however. He argues that there may be a costs associated with futures market participation (perhaps in the form of learning the mechanics of futures market operation), which limit the participation of some types of investors. If this is the case, his theoretical models show that even in the presence of a zero beta, not all risk can be diversified away. He therefore argues that risk premiums in

futures markets could be composed of two components – the standard systematic component, and a “residual” component that is a function of hedging pressure. Bessimbinder (1992) finds that futures returns covary with hedger’s net positions, and concludes that this result supports hedging pressure as a determinant of risk premiums in futures markets, consistent with Hirschleifer’s model, and with the generalized concept of normal backwardation.

We now summarize the above discussion regarding existing empirical research on normal backwardation. Risk premiums that may exist in futures markets cannot be observed, because the expected futures cash price cannot be observed. The standard empirical practice then is to check for speculative profits, which would be consistent with the existence of risk premiums. If speculative profits exist (the evidence on this is mixed), they must be decomposed into profits due to forecasting ability, which is unobserved, and any residual profits (a dubious proposition). If there are profits that are not due to forecasting ability, it is inferred that risk premiums are present. These risk premiums may be due to systematic risk if futures price changes are correlated with returns to total wealth (a nebulous concept). After adjusting “observed” risk premiums for systematic risk, it is then inferred that any residual risk premium that is not due to systematic risk may be due to hedging pressure, if measures of these two phenomena are correlated. This path by which a researcher might find evidence consistent with the generalized theory of normal backwardation is so convoluted, it is little wonder that no consensus has been reached. If such evidence is found and can be believed, it is still

only *consistent* with the theory, the real burning issue of causality is never addressed. Does the net position of hedgers *cause* futures price changes?

We believe that the successful evaluation of this question requires that it be reconsidered from scratch, in a framework that explicitly addresses the issue of causality, and simultaneously accounts for the existence of relevant, but unobserved quantities. We provide such an investigation here. Clearly, the correct answer to Keynes' theory is important for market participants. Should a hedger anticipate losing money on average in exchange for enjoying reduced risk? Can speculators expect to be profitable on average by merely taking a position opposite that of hedgers' net position, regardless of the depth of their knowledge of a market and their forecasting ability?

The second issue that we investigate is the cause(s) of positive correlation between the volume of trade and degree of price variability in futures markets. This relationship is well documented; Karpoff (1987) provides a survey of the evidence. There are two theoretical explanations for this phenomenon. First, there is the Mixture of Distributions Hypothesis (MDH), due originally to Clark (1973). He proposes a model in which there is a stochastic number of independent price changes over any time period, due to a non-constant rate of information arrival. This results in the variance of the overall price change for a given period being an increasing function of number of within-period price changes. Volume of trade is also specified as an increasing function of the number of within-period price changes. Thus, in this theory the (unobserved) rate of information arrival is a common cause of trading volume and price change volatility. Epps and Epps (1976) present an alternative formulation of the MDH. They specify an

equilibrium model of intraday price determination in which the level of disagreement among traders causes the magnitude of a day's overall price change. Here volume is also an increasing function of disagreement, and the Epps & Epps model therefore also implies that volume and volatility are both effects of a latent common cause. Tauchen and Pitts (1983) offer a MDH model that incorporates elements of both the Clark and Epps and Epps models. On a related issue, the MDH models also result in a leptokurtotic distribution for observed price changes, which is consistent with empirical evidence. A competing explanation for this phenomenon due to Mandelbrot (1963) is that price changes are drawn from a distribution with infinite variance from the stable Paretian family. Finding evidence supporting the MDH explanation for positive volume and volatility correlation would thus also support one theory of the cause of excess kurtosis in the futures price change distribution.

The competing explanation for the positive correlation between trading volume and price volatility in futures markets is that of noisy rational expectations (NRE). In the NRE model of Shalen (1993), there are two types of traders. Informed traders have private information regarding market values. Uninformed speculators, on the other hand have no private information, and attempt to extract price signals from observed futures price changes. The series of these price changes is noisy, however, due to a random liquidity demand from hedgers (buying or selling due to their activity in the underlying market, not due to information arrival). The uninformed speculators can then misinterpret this liquidity trading as being due to information arrival, causing them to

adjust their positions, resulting in increases in volume and volatility. The level of activity of uninformed speculators then is a common cause of volume and volatility.

We investigate the causal mechanisms driving the volume and volatility relationship. In addition to potentially vindicated one of the theoretical explanations given above, understanding these mechanisms is important for market participants, researchers, and regulators. All market participants are obviously impacted by price volatility and market depth, and clearly should be interested in the underlying causes. Researchers will be interested in the correct specification of empirical models, and results inconsistent with existing theories might inspire new ones. Regulators have displayed an interest in curbing excessive levels of price volatility, and the success of such an endeavor would be greatly aided by a deep understanding of its causes.

The third issue that we investigate is allegations that the activities of specific types of traders are causes for the level of price volatility. This is closely related to the volume - volatility issue; as explained above, the NRE expectation model of Shalen predicts that volatility is an increasing function of the number of uninformed speculators. Similarly, in the model of Stein (1987), rational, but imperfectly informed futures speculators can (but do not necessarily) destabilize prices. These models contrast with the rational expectations model of Danthine (1978), in which imperfectly informed speculators stabilize prices. More recently, the finance literature has become interested in irrational behavior. An example of this is the model of DeLong, et al. (1990). In their model, irrational traders drive an asset's price away from the fundamental value, Rational arbitrageurs' fear that the return to fundamental value may be slow coming, and

so limit their activity, resulting in increased price volatility. This model is not concerned with futures markets as such, but the underlying principals might still apply.

Empirical evidence compiled regarding this question thus far is limited. Daigler and Wiley (1999) examine various financial futures markets, and report that the activity of futures traders who are on the trading floor is associated with decreased price volatility, while the activity of the “general public” is associated with increased volatility. The on-floor traders can observe the identities of those making large trades, and are therefore in a position to infer the informational content of those trades. They can therefore be thought of as informed, and the results are thus consistent with the models of Shalen and Stein. Chang, Chou, and Nelling (2000) find that in the S&P 500 futures market, large hedging activity is positively correlated with volatility, and concludes that increased volatility likely results in increased hedging demand. Wang (2002) finds that in exchange rate futures markets measures of speculative activity and volatility are positively related. He suggests then that speculators destabilize markets. Note that these last two studies find very similar empirical evidence, but reach opposite conclusions regarding the likely direction of causality. Neither seems to consider the possibility that the observed relationship might be due to a common cause. This demonstrates the merit of inferring causal relationships from observational data using appropriate theory rather than intuition. This question is important for reasons similar to those given above for our second line of inquiry.

INFERRING CAUSAL RELATIONSHIPS FROM INCOMPLETE, OBSERVATIONAL DATA

As mentioned in the introduction, treatments of the theory of causal inference using observational data can be found in Pearl (2000) and Spirtes, Glymour and Scheines (2000). These methods are just beginning to be adopted in applied economic research, although these efforts to date have largely worked under an assumption of causal sufficiency (i.e. that the researcher has collected observations for all variables present in the unknown causal structure). Swanson and Granger (1997) search for causal relationships among the variables in a vector autoregression to guide an appropriate Bernanke decomposition of the innovation covariance matrix and Demiralp and Hoover (2003) investigate the reliability of such a procedure. Haigh and Bessler (2003) investigate price discovery in cash grain markets and a related transportation market. Akleman, et al (1999) investigate causal relationships among corn exports and exchange rates using causal methods, both with and without the assumption of causal sufficiency.

We now provide a description of the algorithm that we employ to inferring causal relationships, the Fast Causal Inference (FCI) algorithm. The FCI was developed to be appropriate for inferring causal relationships from observational data (to the extent possible), even in the presence of latent variables. This section is adapted from Chapter 6 of Spirtes, Glymour and Scheines (2000); see that work for a more thorough description. The causal literature has developed the directed graph as a tool for visually representing a group of related causal relationships. A graph is a set of variables (V_1, V_2, \dots, V_n) that are connected by lines called edges, which may represent causal

flows. If two variables are connected by an edge, they are said to be adjacent. Directed edges have arrowheads on the ends indicating the direction of causal flow between two adjacent variables. For example, $V_1 \rightarrow V_2$ indicates that V_1 is a cause of V_2 . V_1 is a parent of V_2 if there is a directed edge from V_1 to V_2 . A path is a sequence of variables such that each pair of variables that are adjacent in the sequence are also adjacent in the graph. A directed path is a path containing only directed edges in which causal flow runs from the first endpoint on the path to the last. An undirected path is a path in which causal flow is not required to run from the first endpoint on the path to the last. If there is a directed path from V_1 to V_2 , we say that V_1 is an ancestor of V_2 (e.g. as is the case in the graph $V_1 \rightarrow V_3 \rightarrow V_2$) and that V_2 is a descendant of V_1 . Note that parents are always ancestors, but the reverse is not true. A cyclic path is one in which causal flow begins at a variable and eventually returns to that variable (e.g. $V_1 \rightarrow V_2 \rightarrow V_3 \rightarrow V_1$). If a variable is caused by two other variables on a path, it is said to be a collider. For example, in the graph $V_1 \rightarrow V_2 \leftarrow V_3$, V_2 is a collider on the paths $\langle V_1, V_2, V_3 \rangle$ and $\langle V_3, V_2, V_1 \rangle$. A graph that contains directed edges, and no cyclic paths is a directed acyclic graph (DAG). The set of variables in a DAG is assumed to be causally sufficient – there are no latent common causes for any pair of variables in V .

Nature may choose to hide some variables, however. Suppose there is a DAG G over a set of variables V , and that O is a subset of the variables in V that are observed. A path U is an inducing path relative to O if and only if a) every member of O on U , except for the endpoints, is a collider on U , and b) every collider on U is an ancestor of either one of the endpoints. For example, in the graph G (see figure 3.1), the path $U =$

$\langle V_1, V_2, V_3, V_4, V_5, V_6 \rangle$ is an inducing path from V_1 to V_6 over $O = \{V_1, V_2, V_4, V_6\}$. As required, each of the colliders on U , V_2 and V_4 , is an ancestor one of the endpoints, V_1 and V_6 , and the variables on U that are in O (other than the endpoints) are each colliders on U . Inducing paths provide the critical connection between statistical independence relations and causal mechanisms represented in graphs over observable variables. The existence of an inducing path between V_1 and V_2 is implied if V_1 and V_2 are statistically dependent conditional on every subset of $O \setminus \{V_1, V_2\}$ (although this fact alone does not imply the direction of causal flow). This implies that for our example, we would be able to find no subset of $\{V_2, V_4\}$ (including the empty set) that could be conditioned on to render V_1 and V_6 independent.

A graph is an inducing path graph (IPG) over O if there is an edge between two variables V_1 and V_2 with an arrowhead at V_2 if and only if there is an inducing path in G from V_1 to V_2 relative to O . To continue the example from the above paragraph, suppose we observe only the variables in $O = \{V_1, V_2, V_4, V_6\}$. As previously established, there is an inducing path running from V_1 to V_6 over O . The inducing path graph G' over O (shown in figure 3.2) thus features a directed edge running from V_1 and V_6 . Note, however, that in G there was no edge running from V_1 to V_6 . This illustrates an important point – the existence of a directed edge between two variables in an IPG implies that one variable is an ancestor of the other in the underlying DAG, but not necessarily a parent. Note also that some edges in G' have arrowheads on both ends. These result from the existence of inducing paths running from each variable to the other. For example, $\langle V_2, V_3, V_4 \rangle$ is an inducing path over O , as is $\langle V_4, V_3, V_2 \rangle$. Hence

the definition of IPG above requires arrowheads at both ends of the edge between V_2 and V_4 in G' . Such edges are referred to as bidirected edges, and they imply that the two adjacent variables have a latent common cause (in this case V_3).

Unfortunately, the statistical conditional independence relations over a set of observed variables will not necessarily identify a unique IPG. For an IPG G' , the set of all IPGs that entail equivalent sets of statistical independence relations over a given set of observed variables O is denoted $\text{Equiv}(G')$. A partially-oriented inducing path graph (POIPG) is a pattern that represents set of IPGs in $\text{Equiv}(G')$, where G' is the true IPG over O for the DAG G . The ends of the edges in a POIPG can have any one of three types of marks: no mark, and arrowhead, or an “o”. We use the symbol “*” to denote any one of these three types of end marks. We say that π is a POIPG of DAG G with IPG G' over O if and only if: a) π and G' have the same variables and adjacencies; b) if $V_1 \circ \rightarrow V_2$ is in π , then either $V_1 \rightarrow V_2$ or $V_1 \leftrightarrow V_2$ is in every IPG in $\text{Equiv}(G')$; c) if $V_1 \rightarrow V_2$ is in π , then $V_1 \rightarrow V_2$ is in every IPG in $\text{Equiv}(G')$; d) if $V_1 * \text{---} \underline{\underline{V_2}} * \text{---} V_3$ is in π , then V_2 is a non-collider in every IPG in $\text{Equiv}(G')$; e) if $V_1 \leftrightarrow V_2$ is in π , then $V_1 \leftrightarrow V_2$ is in every IPG in $\text{Equiv}(G')$; and f) if $V_1 \circ \text{---} V_2$ is in π , then either $V_1 \rightarrow V_2$, $V_1 \leftarrow V_2$, or $V_1 \leftrightarrow V_2$ is in every IPG in $\text{Equiv}(G')$. The adjacencies that exist in a POIPG then convey information about the conditional independence relations among the observed variables, and end marks on the edges other than “o” convey information about the direction of causal flow in the underlying DAG. The output of the FCI algorithm that we describe below is a POIPG.

One special type of path that may be found in a POIPG is a definite discriminating path, the existence of which may be used when orienting the edges in the FCI algorithm. A path U is a definite discriminating path for V_1 if and only if U is an undirected path between V_2 and V_3 containing V_1 , every variable on U , except the endpoints, is either a collider or definite non-collider on U and the following conditions also hold:

A) If V_4 and V_5 are adjacent on U and V_5 is between V_4 and V_1 on U , then

$$V_4^* \rightarrow V_5$$

B) If V_4 is between V_3 and V_1 on U and V_4 is a collider on U then either

$$V_4^* \rightarrow V_3 \text{ or } V_4 \leftarrow V_3.$$

C) If V_4 is between V_2 and V_1 on U and V_4 is a collider on U then either

$$V_4^* \rightarrow V_2 \text{ or } V_4 \leftarrow V_2.$$

D) V_2 and V_3 are not adjacent.

Some conditions regarding the underlying DAG G are necessary for inferring the set of IPGs over O that are in $\text{Equiv}(G')$ using a set of conditional independence relationships. First, the Markov condition assumes that in the probability distribution over the variables V in the underlying DAG G , a variable V_1 is independent of every set of variables that does not contain V_1 or its decedents, conditional on V_1 's parents. This essentially states that it is possible to represent a set of conditional independence relations graphically, using the definitions of a DAG and the related terminology that we laid out in the first paragraph of this section. Second, the faithfulness or stability condition requires that the conditional independence relations among the variables V in

the underlying DAG G are due to the topology of G , rather than peculiar, offsetting parameter values in the causal relationships. Pearl (2000) gives the following example. Suppose we have DAG H (see figure 3.3), and that the causal relationships are represented by the structural equations

$$V_2 = cV_1 + u_2$$

and

$$V_3 = aV_1 + bV_2 + u_3$$

where u_2 and u_3 are independent stochastic errors. Note that generally we would expect V_2 and V_3 to be dependent, but if the parameter a just happened to take the value $-bc$, then V_2 and V_3 would be independent. The faithfulness or stability condition states that one is unlikely to encounter this type of independence relation in practice. If, in examining a hypothetical data set associated with figure 3.3, the only conditional independency we find is that V_2 and V_3 are independent conditioned on the null set, the faithfulness condition allows us to infer that V_2 and V_3 should not be adjacent, and that neither V_2 nor V_3 is caused by V_1 .

We now describe the Causal Inference (CI) algorithm, the basic functioning of which underlies the FCI algorithm that we use, and then describe how the two algorithms differ. The input of either algorithm is observations over a set of possibly causally insufficient variables, and the output of either algorithm is a POIPG. Both algorithms consist of two phases, determining the adjacencies in the POIPG using a statistical test of conditional independence relationships (Fisher's z test), and then deducing the maximally informative orientation of the resulting edges that is consistent

with the faithfulness condition and the assumption that the underlying graph is a DAG (i.e. there are no cycles). The CI algorithm involves the following steps:

- A) Form a complete undirected graph on the set of variables O , in which every variable is connected to every other variable by an undirected edge.
- B) If two variables V_1 and V_2 are independent conditional on any subset S of $O \setminus \{V_1, V_2\}$, remove the edge between V_1 and V_2 , and record S in the separating set for V_1 and V_2 , denoted $\text{Sepset}(V_1, V_2)$.
- C) Let F be the graph that results from step B). Orient each edge as $o \rightarrow o$. For each triple of variables (V_1, V_2, V_3) such that the pairs (V_1, V_2) and (V_2, V_3) are adjacent in F but the pair (V_1, V_3) is not, orient $V_1^* -^* V_2^* -^* V_3$ as $V_1^* \rightarrow V_2^* \leftarrow^* V_3$ if and only if V_2 is not in $\text{Sepset}(V_1, V_3)$ and arrange $V_1^* -^* V_2^* -^* V_3$ as $V_1^* -^* \underline{V_2^*} -^* V_3$ if and only if V_2 is in $\text{Sepset}(V_1, V_3)$.
- D) Repeat the following sequence of instructions until no more edges can be oriented:
 - i) If there is a directed path from V_1 to V_2 , and there is an edge $V_1^* -^* V_2$, orient $V_1^* -^* V_2$ as $V_1^* \rightarrow V_2$.
 - ii) Else, if V_1, V_2, V_3 is a collider along $\langle V_1, V_2, V_3 \rangle$, V_1, V_2, V_3 is adjacent to V_1, V_2, V_3 , and V_1, V_2, V_3 is not in $\text{Sepset}(V_1, V_3)$, then orient $V_1^* -^* V_2$ as $V_1 \leftarrow^* V_2$.
 - iii) Else, if U is a definite discriminating path between V_1 and V_2 for V_3 , and V_4 and V_5 are adjacent to V_3 on U , and V_3, V_4 , and V_5 form a triangle, then

- a) If V_3 is in $\text{Sepset}(V_1, V_2)$, then mark $V_3^* \text{---}^* V_4^* \text{---}^* V_5$ as $V_3^* \text{---}^* \underline{V_4^*} \text{---}^* V_5$
- b) Else, $V_3^* \text{---}^* V_4^* \text{---}^* V_5$ as $V_3^* \rightarrow V_4^* \leftarrow^* V_5$
- iv) Else, if $V_1^* \rightarrow \underline{V_2^*} \text{---}^* V_3$ then orient as $V_1^* \rightarrow V_2^* \rightarrow V_3$.

Step B above is computationally infeasible, as the number of possible subsets of O grows very rapidly with the cardinality of O . Checking for conditional dependence of two variables V_1 and V_2 over all possible subsets of $O \setminus \{V_1, V_2\}$ then becomes very difficult. The FCI and CI algorithms differ in the way that step B is performed. The FCI uses an intermediate step to infer that some variables cannot be in $\text{Sepset}(V_1, V_2)$, thereby reducing the number of conditional independence tests that must be performed. This procedure is relatively complicated, and does not offer any additional understanding of the means by which the causal structure is inferred, and we therefore do not describe it. The important fact is that the FCI algorithm is essentially a computationally feasible version of the CI algorithm. The FCI algorithm is implemented in the Tetrad 3 computer program, which we use in our analysis.

DATA

We analyze eight futures markets: Chicago Board of Trade (CBOT) corn, New York Mercantile Exchange (NYMEX) crude oil, Chicago Mercantile Exchange (CME) Eurodollar deposits, New York Commodity Exchange (COMEX) gold, CME Japanese Yen, New York Board of Trade (NYBOT) coffee, CME live cattle, and the CME S&P 500. Observations for all data over the interval March 21, 1995, through January 8,

2003, are used. We construct three types of data series for use in the analysis: i) those related to trader activity and positions, ii) those related to futures prices and trading volume, and iii) trend and seasonal series. We now discuss each category of data in turn.

The Commodity Futures Trading Commission (CFTC) requires certain exchange members and futures commission merchants (i.e. brokers) to file daily reports with the Commission. Those reports show the futures positions of traders that hold positions above specific reporting levels set by CFTC regulations (these are referred to as “reportable positions”). Each trader is classified as being either commercial or non-commercial, with commercial traders being those engaged in hedging activity. Ederington and Lee (2002) caution that this distinction is not always entirely accurate, and our data regarding trader type are thus noisy. Henceforth we refer to reportable commercial positions as being those of “large hedgers”, to reportable non-commercial positions as being those of “large speculators”, and to non-reportable positions as being those of “small traders”. The data collected as of a markets close on each Tuesday are released to the public in the CFTC’s Commitments of Traders (COT) report, generally on the following Friday. We use this data in two ways. First, we calculate the net position of large hedgers (*LH Net Position*) as the number of open long futures positions minus the number of open short futures positions held by large hedgers. Second, we calculate the aggregate level of activity of each trader type (*LH Activity*, *LS Activity*, and *ST Activity* for large hedgers, large speculators and small traders, respectively) as the sum of their open long and short futures positions. These three variables, at any point in time, sum to twice the level open interest in the market. Some adjustments to the COT

data are necessary. Before 1998, corn futures positions are measured in numbers of 1,000 bushels, rather than number of contracts (each calling for delivery of 5,000 bushels). We therefore divide all corn COT data prior to 1998 by five so that the related data series we use are measured in consistent units over the sample period. The size of the cash settlement called for by the S&P 500 futures contract was halved in late 1997, and we therefore multiply all S&P 500 COT data prior to the change by two, to make our measures of trader positions consistent with the current contract specification. In the crude oil and coffee markets, observations for the COT series are missing for September 11, 2001, and are linearly interpolated.

Daily price data for individual deliveries for each market are provided by Commodity Research Bureau, and the corresponding volume data are provided by Primark Datastream. We construct a continuous futures price level series (*Nearby*) for each market using week-ending observations of the futures contract nearest to expiration. We use weeks that run Wednesday through Tuesday in constructing all price, volume, and volatility series, so as to correspond with the COT data. We also construct a nearby weekly returns series (*Return*) using weekly returns series for each individual delivery. Thus no observations in our *Return* series are constructed using price level observations from two different deliveries (as would be the case if one simply constructed a return series using a previously constructed nearby levels series). A weekly total volume series (*Volume*) was constructed for each market by summing the total trading volume for all deliveries for each day. A measure of futures price volatility

(*Volatility*) was constructed by taking the log difference between the high and low prices for each week for the nearby contract.

A linear time trend series (*Time*) is used in the analysis, as are weekly observations of two annual seasonal harmonic variables. These are defined as

$$Annual\ Sin = \sin\left(\frac{2\pi Time}{52}\right)$$

and

$$Annual\ Cos = \cos\left(\frac{2\pi Time}{52}\right).$$

These harmonic series account for the possibility of seasonal influences on the volume, volatility, and activity variables which we expect in the agricultural commodity futures markets.

ANALYSIS

We begin by investigating the possibility of non-stationary behavior in the series. *A priori*, the Efficient Market Hypothesis gives us strong reason to suspect that the *Nearby* series may contain a unit root, however, we have no such grounds for suspicion with respect to the remaining series. Indeed, it would seem rather implausible to believe that *LH Net Position*, for example, might drift off toward infinity. All data series save the trend and seasonal harmonic series are subjected to Augmented Dickey-Fuller (ADF) tests for non-stationary, with the results given in table 3.1. We find that for seven of the eight *Nearby* series we cannot reject the null hypothesis of no unit-root, confirming our initial suspicion. We therefore use the *Return* series for all markets in the causal

analysis that follows (we use the *Return* series even the live cattle market, as we wish to keep the interpretation of the results consistent across markets). For the remaining series (*Volatility*, *Volume*, *LH Activity*, *LS Activity*, *ST Activity*, and *LH Net Position*), we generally reject the null hypothesis that each series contains a unit root. Given that a) the burden of proof was on proving that there is no unit-root, b) our prior expectations, c) it is a well-established fact that ADF tests have low power against plausible alternatives (see, for example, DeJong, et al, 1992), and d) our desire to use consistent types of series (i.e. levels or differences) across markets, we proceed to use the levels series for all variables other than *Return*.

We apply the FCI algorithm to the 10 data series for each market. In all cases, the algorithm is restricted from allowing inducing paths running from any variable to *Time*, and from allowing a latent common cause for any variable and *Time*. Similar restrictions are placed on the allowed orientations of edges attached to the seasonal harmonic series, although the possibility of inducing paths running from *Time* to either of the seasonal harmonic series is not prohibited. The resulting POIPGs are presented in figures 3.4 through 3.11.

We first consider the evidence with regard to the generalized theory of normal backwardation. Our analysis considers the relationship between week-ending level of *LH Net Position* and the *Return* that was realized over those weeks. We interpret this as follows. Suppose the futures price begins at exactly the unobserved spot price that is expected to prevail at the time of expiration. A move by *LH Net Hedging* to a higher level should, if normal backwardation holds, then cause the futures price to move higher,

to some price above the expected future spot price so that speculators who are now more short can be compensated. We then expect a positive relationship between the week-ending level of *LH Net Hedging* and *Return*. The *prima facie* evidence in this regard is not generally supportive of the hypothesis that hedging pressure causes risk premiums, as we find a negative correlation between these two variables in all markets except the S&P 500. Examining the POIPGs for the eight markets, we find that *LH Net Hedging* and *Return* have a latent common cause in three markets (gold, japanese yen, and coffee), no causal connection in three markets (Eurodollar deposits, live cattle, and the S&P 500), and that in the remaining two markets (corn and crude oil) either there is a latent common cause or there is an inducing path from running *Return* to *LH Net Position* in all IPGs consistent with the observed set of conditional independencies. In no case do we find the possibility that causal flow might run from *LH Net Hedging* to *Return*, and we firmly conclude that hedging pressure does not cause returns, and we thus find no support for the generalized theory of normal backwardation. We can conclude, then, that it does not appear that hedgers need not expect to automatically pay a risk premium to speculators. Note, however, that this is not the same as concluding that risk premiums do not exist in these markets, only that there are not risk premiums *caused* by hedging pressure. The speculative profits sometimes found in other research could then be due to speculators collecting risk premiums that are due to other causes, or due to superior forecasting ability.

We now describe how the algorithm arrived at this conclusion. In the cases where the two variables are adjacent in the POIPG, a sufficient condition to conclude

that causal flow does not run from *LH Net Hedging* to *Return* is the existence of an arrowhead on the *LH Net Hedging* end of the edge. We explain the existence of such an arrowhead using the corn market as an example. After the adjacencies are determined for the corn market POIPG (step B in the FCI algorithm), the following sub-graph is present: $Volume \circ LH\ Net\ Hedging \circ Return$. The adjacency between *Volume* and *Return* is removed because the unconditional correlation between the two is not significantly different from zero. Finding that *Volume* and *Return* are unconditionally uncorrelated prevents us from believing that we could have either $Volume \rightarrow LH\ Net\ Hedging \rightarrow Return$ or $Volume \leftarrow LH\ Net\ Hedging \leftarrow Return$. Furthermore, the faithfulness condition prevents us from believing that $Volume \leftarrow LH\ Net\ Hedging \rightarrow Return$ (if this were the case, it would be very unusual to find that *Volume* and *Return* were unconditionally uncorrelated). We therefore must accept the only remaining possibility, that $Volume \rightarrow LH\ Net\ Hedging \leftarrow Return$ is the appropriate orientation (*LH Net Hedging* is a collider). We thus have an arrowhead at the *LH Net Hedging* end of the edge between it and *Return*. This type of edge orientation is due to step C of the FCI algorithm, and this rule can be used to orient all of the edge end marks that are critical to our analysis in this chapter.

We next discuss the evidence regarding relationships between trader type and volatility levels, and afterwards discuss the related issue of the *Volume* and *Volatility* relationship. The theories discussed earlier in the chapter make predictions regarding causal relationships between speculators and/or uninformed traders. *LS Activity* obviously represents speculative activity, and some would argue that this category

represents uninformed traders to some extent as well. It does not seem unreasonable to interpret *ST Activity* as representing uninformed traders. Such distinctions turn out not to be necessary, however. We find no evidence of causal flow running from either *LS Activity* or *ST Activity* to *Volatility* in any of the eight markets. The edges directly connecting *ST Activity* and *Volatility* are removed by conditioning on the empty set in five markets, by conditioning on *Volume* in crude oil and eurodollars, and by conditioning on the *Time* trend for the S&P 500 market. The edge between *LS Activity* and *Volatility* is removed in the following markets by conditioning on the variables given in parentheses: corn (*Annual Sin*), crude oil (*LH Activity*), Eurodollars (*Return*), gold (*Volume* and *Time*), japanese yen and coffee (the empty set), live cattle (*LH Net Position*), and S&P 500 (*Time*). This information is summarized in table 3.2. Two variables need not be connected directly by an inducing path in the POIPG for causal flow to run between them – we may find a roundabout directed path from one variable to the other. We find no evidence of such indirect causal flow in this case however. We thus find no evidence supporting theories that predict that the activity levels of speculators and/or uninformed traders affects volatility (either positively or negatively).

With regard to the *Volume* and *Volatility* relation, we find that in six of the eight markets there is a latent common cause for the two variables. In the coffee market, there is either a latent common cause, or that there is an inducing path from *Volatility* to *Volume* in all observationally equivalent IPGs. The evidence from these markets is then consistent with the MDH, which predicts that either the rate of information arrival (the Clark version) or the level of disagreement among traders (the Epps and Epps version)

or some combination of these causes the positive correlation. In the crude oil market, we find an inducing path running from *Volume* to *Volatility*, which is consistent with neither the MDH nor Shalen's prediction that the level of activity of informed speculators causes the positive correlation between *Volume* and *Volatility*. If Shalen's theory is true, we expect to find causal flow running from either *LS Activity* or *ST Activity* to both *Volume* and *Volatility*. We find no such evidence in any of the markets that we analyze. The edges between the activity levels and *Volatility* were removed for the reasons discussed previously. The edges between *Volume* and *ST Activity* generally are not removed (crude oil and Eurodollars being the exceptions), but are bidirected, implying a latent common cause. Edges between *LS Activity* and *Volume* are removed in all markets save one (live cattle). This was accomplished by conditioning on *LH Activity* (Eurodollars, Japanese yen, and coffee), *Volatility* (gold and S&P 500), and the empty set (corn). Thus most of the necessary edges to support Shalen's theory are removed, and even though edges connecting *ST Activity* and *Volume* are generally not removed, they are bidirected implying a latent common cause. We also again find no evidence of indirect causal paths that would support Shalen's theory.

The practical implication of our findings is that those attempting to model time-varying volatility may indeed find volume to be a useful proxy for some unobserved cause or causes of volatility. It is not necessarily prudent, however, to assume that any event that will affect an increase in volume will also result in an increase in volatility. Contract expiration, for example, generally results in increased volume as traders roll positions out of the expiring contract. This event has nothing to do with either of the

unobservable common causes of volume and volatility that have been suggested in the theoretical literature, and should not therefore be expected to cause an increase in volatility.

CONCLUSIONS

In this article, we examine various unresolved issues regarding causal relationships in futures markets. To this end, we apply the Fast Causal Inference (FCI) algorithm, which has been developed in the formal causality literature as an appropriate tool for inferring causal relationships using observational data, even in the presence of relevant unobserved quantities. Such an approach is highly attractive, considering that most research in empirical economics and finance is conducted in such an environment. We find no support for the generalized theory of normal backwardation, and thus no reason to believe that hedgers will generally transfer a risk premium to speculators in exchange for risk-bearing services. We find no support for theories predicting that particular types of traders affect the level of price volatility, either positively or negatively, in futures markets. We find evidence that supports the mixture of distributions hypotheses (MDH), which posit the existence of one or more unobservable common causes of trading volume and price volatility. This suggests that models of time-varying volatility can benefit from the information about the latent variable(s) contained in volume, but caution in the interpretation of such a model is necessary as volume does not actually cause volatility.

There are abundant opportunities for the further application of causal inference methods to empirical research into derivatives markets. Other open questions need to be addressed, some of which are: is the level of futures trading activity a cause of price volatility in the underlying cash market? What are the causes and/or effects of changes in the shape of the forward curve? What are the causes of basis movements? Does the size of the margin deposit required to trade futures impact any of the quantities that we have considered? What are the causal relationships that exist across related markets (e.g. the soy complex or the crude oil complex)? Although the theoretically correct procedures for doing so are not obvious at this time, it would be very interesting to investigate the causal relationships that exist among the variables that we considered in a dynamic setting (e.g. does $Volatility_{t-1}$ cause $Volatility_t$, or is there a latent common cause for both?). These issues offer a fertile ground for future research.

CHAPTER IV

COMPARING THE PERFORMANCES OF THE PARTIAL EQUILIBRIUM AND TIME SERIES APPROACHES TO HEDGING

INTRODUCTION

Two broad strategies for optimally hedging risky market commitments have emerged in the academic literature and in practice. The applied economics literature has focused on the use of statistical models of the observed time series of cash and futures prices in hedging. Early development of this type of optimal hedging is found in Johnson (1960), Peck (1975), and Kahl (1983), among others. Typically, this type of hedging considers an agent with a non-tradable position in a cash commodity, who plans to buy or sell some number of commodity futures contracts that will maximize her utility. This traditionally involved making static estimates of the variances of changes in the cash and futures prices and the covariance between those changes, and then choosing a level of hedging that would minimize the variance of changes in the hedger's portfolio value. More recently, Cecchetti, Cumby and Figlewski (1988), Myers (1991), and Baillie and Myers (1991) have adopted the use of models of time-varying conditional variance for optimal hedging. Noting that the use of differenced data will lose information about the long-run relationship between two time series, Kroner and Sultan (1993) incorporate the co-integrating relationship between cash and futures prices into their model. Gagnon, Lypny and McCurdy (1998) and Haigh and Holt (2000) extend these models to include multiple risks.

The finance literature on the other hand has typically stressed the use of partial equilibrium derivative pricing models for hedging. This began when Black and Scholes (1973) and Merton (1973) noted that the seller of a derivative could form a risk-free portfolio by holding just the right quantity of the underlying security. This quantity is determined by the rate at which the price of the derivative will change as the price of the underlying changes, referred to as the “delta.” This type of hedging is therefore often referred to as “delta hedging.”

In application, different types of hedgers have tended to make use of the two strategies. Holders of large derivative portfolios generally have employed partial equilibrium hedging. This is the realm of financial institutions and “financial engineers” that sell derivatives to their customers at a markup to the value of a portfolio with price dynamics that replicate, as closely as possible, those of the derivative. Commodity producers and consumers, on the other hand, have more often used the time series approach. A typical picture is that of the agricultural producer with a crop in the ground, who wishes to minimize the risk that the price of the output will fall before the harvest. Despite their differences, these two types of hedgers face exactly the same problem: they each hold a position in one market (either underlying or derivative), and wish to take a position in the other market that will result in maximum benefit. Either hedger might use either of the two approaches to hedging, despite the traditional divide.

Each approach has its own merits and drawbacks. The time series approach does not require the imposition of theory *a priori*, thereby avoiding potential misspecification. Also, available time series models can very effectively represent time-varying

covariability among price series, a commonly observed market phenomenon that is central to the hedging problem. This approach does not, however, make use of all available information. For example, time series hedging models consider neither the arbitrage activity that constrains the price of a derivative relative to its underlying security, nor theories regarding derivatives' price variability (e.g. the Samuelson (1965) hypothesis that a futures contract's volatility should increase as expiration approaches).

By contrast, the partial equilibrium approach directly incorporates the arbitrage relationship(s) between the derivative and underlying instrument(s). An additional benefit of this approach is the ability to use observed market prices to infer the expectations of market participants. For example, option prices can be used to infer the future levels of volatility that knowledgeable industry participants are anticipating in an associated underlying market. The adoption of the partial equilibrium approach comes at the price however of requiring various simplifying assumptions, which have varying degrees of implausibility. Crucially, most partial equilibrium models do not incorporate the stochastically time-varying volatility that is widely acknowledged to exist in most financial and commodity markets.

Given the above stated benefits and drawbacks of each of the two approaches to hedging, it is not immediately clear that one approach should be preferred in any given situation. No previous research has directly compared the effectiveness of these two hedging strategies, and we thus undertake such an evaluation here. We directly compare the in-sample and out-of-sample hedging performance of the two approaches for a trader that is long physical crude oil, and uses a simple derivative with a linear payoff function

(a futures contract) to hedge the associated price risk. We assume that the hedger maximizes a mean-variance utility function, and hedging effectiveness is measured by the increases in utility that the hypothetical trader realizes by implementing each strategy (relative to not hedging at all). Two partial equilibrium models that have been developed in the commodity contingent claims pricing literature are considered – the Schwartz (1997) one-factor model, and the two-factor model of Gibson and Schwartz (1990). Various strategies for estimating and inferring these models' parameters are employed. The competing time series model is a vector error-correction model, with a generalized autoregressive conditional heteroskedastic error structure.

The rest of this chapter is organized as follows: In the following section, we describe the hedging problem and the time series model. The next section describes the partial equilibrium models, and describe how they can adapted for optimal hedging by an agent with mean-variance type utility. This continuous time mean-variance hedging can be considered a generalization of delta hedging. We also show these models can be extended to allow for spatial and form differences between the commodity to which a hedger is committed and the commodity underlying the futures contract. We then discuss the estimation of all models, and report their hedging effectiveness, before offering some concluding remarks.

HEDGING COMMODITY PRICE RISK USING TIME SERIES MODELS

We consider a hedger that is long a physical commodity, and wishes to optimally select of a quantity of futures contracts to sell. The hedge ratio, b , is the ratio

of the size of the futures market position to the size of the cash market position. The change in the hedger's portfolio value over the discrete interval from time $t-1$ to time t is given by

$$P_t - P_{t-1} = (L_t - L_{t-1}) - b_{t-1}(F_t - F_{t-1}) \quad 4.1$$

where P_t , L_t , and F_t represent portfolio value, the local cash price of the commodity held by the hedger, and the futures price, respectively, in period t . Note that the commodity held by hedger does not necessarily correspond exactly to the commodity underlying the futures contract. The hedger may be holding a different grade of the commodity than is called for by the futures contract (or a different commodity altogether), and she may not be able to deliver her commodity against the futures contract at par value locally. We therefore distinguish between a local cash price of an arbitrary commodity, and the price at the specified futures delivery location of the specified commodity. We refer to the former as a local cash price L_t as above, and to the latter as the spot price S_t .

We assume that the hedger maximizes mean-variance type objective. This is equivalent to maximizing constant relative risk aversion utility when end-of-period terminal wealth is normally distributed (Hey 1979). Furthermore, under such circumstances the mean variance objective given below is the expected certainty equivalent income. The hedger's problem for each period is formulated as follows:

$$\text{Max}_{b_{t-1}} \left[E(\Delta P_t \mid \Omega_{t-1}) - \frac{\lambda_U}{2} \text{var}(\Delta P_t \mid \Omega_{t-1}) \right] \quad 4.2$$

where E is the conditional expectation operator, ΔP_t is the change in portfolio value from $t-1$ to t , Ω_{t-1} is the information available as of $t-1$, λ_U is the coefficient of absolute risk aversion, and $\text{var}()$ is the conditional variance operator. Note that the risk-minimizing objective is a special case of equation 4.2 where $\lambda_U = \infty$. Note that the conditional variance term in equation 4.2 can be expanded, using equation 4.1, as

$$\text{var}(\Delta L_t | \Omega_{t-1}) + b_{t-1}^2 \text{var}(\Delta F_t | \Omega_{t-1}) - 2b_{t-1} \text{cov}(\Delta L_t, \Delta F_t | \Omega_{t-1}) \quad 4.3$$

where $\text{cov}()$ is the conditional variance operator. The objective-maximizing hedge ratio is then given by

$$b_{t-1} = \frac{-\lambda_U^{-1} E(\Delta F_t | \Omega_{t-1}) + \text{cov}(\Delta L_t, \Delta F_t | \Omega_{t-1})}{\text{var}(\Delta F_t | \Omega_{t-1})}. \quad 4.4$$

The second-order condition for this problem is the negative of the risk aversion coefficient multiplied by the conditional variance of changes in the futures price, and we are thus guaranteed a global maximum for a risk-averse hedger. If we have $\lambda_U = \infty$, the first term in the numerator is zero and we have the minimum-variance hedge ratio. For $0 > \lambda_U > \infty$, the optimal hedge ratio contains the minimum-variance component, and a speculative component. If our hedger anticipates a decrease in the futures price, he will reduce the level of hedging to below the minimum variance level to avoid losses in the futures market. Likewise, an anticipated increase in the futures price will compel our hedger to increase the size of the futures position.

Calculating the optimal hedge ratio in equation 4.4 requires the time-series modeler to provide two types of information – the conditional expected futures price

change and conditional variance-covariance forecasts. Recent academic hedging research advocates obtaining the first piece of information using a vector error correction (VEC) model. This is the appropriate modeling technique in the event that each of the two price series is found to follow a unit root process, but a linear combination of the two is found to be stationary (Engle and Granger, 1987). This linear combination is interpreted as representing a long-run equilibrium between the two levels series. The VEC model is essentially a vector auto-regression model in which a deviation from the long-run equilibrium (the “error”) in one time period is subject to some degree of correction in the following time period. A basic representation of a VEC for 2 variables is as follows:

$$\Delta y_t = \pi_0 + \sum_{i=1}^r \pi_i \Delta y_{t-i} + \alpha \beta y_{t-1} + \varepsilon_t \quad 4.5$$

where y_t is the 2×1 vector of observations at time t , π_0 is a 2×1 parameter vector, π_i is a 2×2 coefficient matrix, β is the co-integrating vector characterizing the long-run equilibrium, α is a 2×1 coefficient vector, and ε_t is a vector of innovations. The inner product βy_{t-1} is the deviation from the long-run equilibrium, and α characterizes the rate at which each of the two variables responds to this deviation. Forming y using cash and futures prices, Equation 4.5 can then be used to generate forecasts of futures price changes – one of the components of the optimal hedge ratio above.

The other pieces of information that are required to calculate the optimal hedge ratio in equation 4.4 are the conditional variances and covariance. These can be forecast using multivariate versions of the auto-regressive conditional heteroskedasticity (ARCH)

model of Engle (1982) or the generalized ARCH (GARCH) model of Bollerslev (1986). A GARCH error structure implies that the conditional second moment of the innovation vector of a model follows an autoregressive, moving average process – it is a function of past innovation vectors and past second moments. Here we employ a GARCH(1,1) model with the diagonal vech parameterization of Bollerslev, Engle, and Wooldridge (1988). The conditional distribution of the error vector from equation 4.5 is then given by

$$\varepsilon_t | \Omega_{t-1} \sim N(0, H_t) \quad 4.6$$

$$\text{vech}(H_t) = W + A \text{vech}(\varepsilon_{t-1} \varepsilon_{t-1}^T) + B \text{vech}(H_{t-1}). \quad 4.7$$

Here, $\text{vech}()$ is the column stacking operator that stacks the lower triangular portion of a symmetric matrix, W is a 3×1 vector of constants, and A and B are a diagonal 3×3 coefficient matrices. Equation 4.7 can be used to form one-period ahead forecasts of the variance of futures price changes and the covariance between futures and cash price changes. The VEC-GARCH model given by equations 4.5 through 4.7 thus provides a means by which a hedger can select the optimal level of hedging.

HEDGING COMMODITY PRICE RISK USING PARTIAL EQUILIBRIUM MODELS

Early models for pricing contingent claims included only a single stochastic factor, the price of the underlying asset. These models assumed that a risk-free portfolio consisting of a short position in the derivative contract and a long position in the underlying asset could be formed, and that this portfolio should earn the risk-free rate of

return. Ross (1978) noted that this assumption is inappropriate in the event that there are benefits to holding an actual asset, rather than merely holding a contract calling for future delivery. When the asset is a commodity, the flow of these benefits is referred to as a convenience yield. Kaldor (1939) describes this phenomenon, and it features prominently in the theory of storage developed in Working (1949) and Brennan (1959). Consideration of the convenience yield motivated the development of the Brennan and Schwartz (1985) model for pricing commodity contingent claims, which assumed that a commodity's convenience yield was a constant proportion of the spot price. This assumption that the convenience yield could be specified as a deterministic function of a commodity's spot price was investigated empirically in Brennan (1991), and Gibson and Schwartz (1991). Both studies decisively concluded that such an assumption was inappropriate, and that the convenience yield should be specified as a second stochastic factor.

Gibson and Schwartz (1990) thus developed a model for pricing commodity contingent claims with two stochastic factors, the first being the spot price of the commodity and the second being the instantaneous net (of storage costs) convenience yield of the commodity. In this model, the holder of a commodity derivative faces not only the risk that the spot price of the commodity will change, but also the risk associated with changes in the convenience yield. As it is not possible to hedge the latter risk, the hedger will not be able to form a completely risk-free portfolio, and the Gibson-Schwartz (GS) model is one of incomplete markets.

The GS model assumes that the spot price of a commodity S and associated instantaneous net convenience yield δ follow the joint diffusion process given by

$$dS/S = \mu dt + \sigma_1 dz_1 \quad 4.8$$

$$d\delta = k(\alpha - \delta)dt + \sigma_2 dz_2 \quad 4.9$$

where μ is the drift of spot price returns, σ_1^2 and σ_2^2 are the instantaneous variances of spot price returns and the convenience yield respectively, dz_1 and dz_2 are increments to correlated Brownian motions, with the multiplication rule $dz_1 dz_2 = \rho_{12} dt$, and ρ_{12} being the correlation coefficient. The convenience yield is assumed to revert at rate k to a long-run mean level α . By Ito's Lemma, the price $G(S, \delta, \tau)$ of a commodity contingent claim that is a function of time, and a twice continuously differentiable function of S and δ then follows the diffusion

$$\begin{aligned} dG = & \left[-G_\tau - \frac{1}{2} G_{SS} \sigma_1^2 S^2 + G_{S\delta} S \rho_{12} \sigma_1 \sigma_2 + \frac{1}{2} G_{\delta\delta} \sigma_2^2 + G_S \mu S \right] dt \\ & + [\sigma_1 S G_S] dz_1 + [\sigma_2 G_\delta] dz_2 \end{aligned} \quad 4.10$$

where $\tau = T - t$ is the length of time from the present (t) until expiration of the derivative (T), and G_X represents the partial derivative of G with respect to X . Gibson and Schwartz present a no-arbitrage argument that leads to following partial differential equation that must be satisfied by the price $F(S, \delta, \tau)$ of a futures contract:

$$\frac{1}{2} F_{SS} S^2 \sigma_1^2 + \frac{1}{2} F_{\delta\delta} \sigma_2^2 + F_{S\delta} S \rho_{12} \sigma_1 \sigma_2 + F_\delta S(r - \delta) + F_\delta [k(\alpha - \delta) - \lambda \sigma_2] - F_\tau = 0$$

where r is the risk-free rate of return, and λ is the market price of convenience yield risk. The solution to equation 4.11, as reported in Hilliard and Reis (1998)¹⁰ is

$$F(S, \delta, \tau) = S \exp \left\{ \left[\frac{-\delta(1-e^{-k\tau})}{k} \right] + \tau \left[r - \alpha + \frac{\lambda\sigma_2}{k} + \frac{\sigma_2^2}{2k^2} - \frac{\sigma_1\sigma_2\rho_{12}}{k} \right] \right. \\ \left. + \left(\frac{1-e^{-k\tau}}{k} \right) \left[\alpha - \frac{\lambda\sigma_2}{k} - \frac{\sigma_2^2}{k^2} + \frac{\sigma_1\sigma_2\rho_{12}}{k} \right] + \left[\frac{\sigma_2^2(1-e^{-2k\tau})}{4k^3} \right] \right\}. \quad 4.12$$

We now turn to the task of adapting the GS model for use in hedging. Using equation 4.12 to find the appropriate partial derivatives to substitute into equation 4.10, we find the diffusion followed by a futures contract to be

$$dF = \left[F \left(\mu - (r - \delta) - \frac{\lambda\sigma_2}{k} (1 - e^{-k\tau}) \right) \right] dt + [F\sigma_1]dz_1 + \left[F \left(\frac{-\sigma_2(1 - e^{-k\tau})}{k} \right) \right] dz_2. \quad 4.13$$

For a hedger whose local cash price corresponds to the spot price, changes in portfolio value are given by $dP = dS - b dF$. Using this, applying Ito's Lemma to equation 4.8, and using equation (13), we find that the short hedger's portfolio dynamics are described by the diffusion

$$dP = \left[S \left(\mu + \frac{\sigma_1^2}{2} \right) - bF \left(\mu - (r - \delta) - \frac{\lambda\sigma_2}{k} (1 - e^{-k\tau}) \right) \right] dt \\ + [(S - bF)\sigma_1]dz_1 + \left[bF \left(\frac{\sigma_2(1 - e^{-k\tau})}{k} \right) \right] dz_2 \quad 4.14$$

¹⁰ Gibson and Schwartz (1993) and Schwartz (1997) also publish formulas for the price of a futures contract in the GS model, but these formulas appear to suffer from typographical errors as they do not seem to solve partial differential equation (11).

Defining another standard Brownian motion z and a parameter σ_p such that

$$\sigma_p dz = [(S - bF)\sigma_1]dz_1 + \left[bF \left(\frac{\sigma_2(1 - e^{-k\tau})}{k} \right) \right] dz_2, \quad 4.15$$

we can simplify equation 4.14 to

$$dP = \mu_p dt + \sigma_p dz \quad 4.16$$

with drift $\mu_p = \left[S \left(\mu + \frac{\sigma_1^2}{2} \right) - bF \left(\mu - (r - \delta) - \frac{\lambda \sigma_2}{k} (1 - e^{-k\tau}) \right) \right]$ and instantaneous

variance

$$\begin{aligned} \sigma_p^2 = & S^2 \sigma_1^2 + b^2 F^2 \left[\sigma_1^2 - 2\sigma_1 \sigma_2 \rho_{12} \left(\frac{(1 - e^{-k\tau})}{k} \right) + \sigma_2^2 \left(\frac{(1 - e^{-k\tau})}{k} \right)^2 \right] \\ & - 2bSF \left[\sigma_1^2 - \sigma_1 \left(\frac{\sigma_2(1 - e^{-k\tau})}{k} \right) \rho_{12} \right] \end{aligned} \quad 4.17$$

This expression for the instantaneous variance of changes in portfolio value is analogous to equation 4.3 – the first term is the instantaneous variance of spot price changes, the second term is b^2 multiplied by the instantaneous variance of futures price changes, and the third term is $-2b$ multiplied by the instantaneous covariance between spot and futures price changes. Armed with the above specification for the controlled stochastic process followed by the hedger's portfolio, we are in a position to solve the continuous time version of the hedging problem given by equation 4.2. In the context of the GS model, we find the following expression for the optimal hedge ratio:

$$b_{GS} = \frac{-\lambda_U^{-1} \mu_F + SF \left[\sigma_1^2 - \sigma_1 \left(\frac{\sigma_2 (1 - e^{-k\tau})}{k} \right) \rho_{12} \right]}{F^2 \left[\sigma_1^2 - 2\sigma_1 \sigma_2 \rho_{12} \left(\frac{(1 - e^{-k\tau})}{k} \right) + \sigma_2^2 \left(\frac{(1 - e^{-k\tau})}{k} \right)^2 \right]} \quad 4.18$$

where

$$\mu_F = F \left(\mu - (r - \delta) - \frac{\lambda \sigma_2}{k} (1 - e^{-k\tau}) \right). \quad 4.19$$

Note that the above expression for the optimal hedge ratio has been developed for a hedger whose cash market commitment exactly corresponds to the commodity underlying the futures contract (i.e. $L = S$). This result is of limited usefulness, as many hedgers' cash market commitments vary from the specifications of the futures contract. The GS model can be augmented, however, to derive a more general formulation. We define the difference between the hedger's cash price and the spot price as

$$B \equiv L - S \quad 4.20$$

and we propose the following stochastic process for B :

$$dB = \gamma(\beta - B)dt + \sigma_3 dz_3 \quad 4.21$$

where σ_3^2 is the instantaneous variance of changes in B , dz_3 is a third Brownian motion, and we add the multiplication rules $dz_1 dz_3 = \rho_{13} dt$ and $dz_2 dz_3 = \rho_{23} dt$. We assume that B reverts to level β at rate γ . The mean-reverting nature of B is justified in the event that a stable long-run relationship between the cash and spot prices exists. In the event that no such relationship existed, the futures contract would make an inappropriate hedging vehicle for the cash price concerned. Changes in the hedger's portfolio are then

given by $dP_A = dB + dS - bdF$, and we can follow a succession of steps similar to those above to arrive at the following diffusion:

$$dP_A = \mu_{AP}dt + \sigma_{AP}dz \quad 4.22$$

with drift $\mu_{AP} = \left[\gamma(\beta - B) + S \left(\mu + \frac{\sigma_1^2}{2} \right) - bF \left(\mu - (r - \delta) - \frac{\lambda \sigma_2}{k} (1 - e^{-k\tau}) \right) \right]$ and

instantaneous variance

$$\begin{aligned} \sigma_{AP}^2 = & \left[S^2 \sigma_1^2 + 2S\sigma_1\sigma_3\rho_{13} + \sigma_3^2 \right] \\ & + b^2 F^2 \left[\sigma_1^2 - 2\sigma_1\sigma_2\rho_{12} \left(\frac{(1 - e^{-k\tau})}{k} \right) + \sigma_2^2 \left(\frac{(1 - e^{-k\tau})}{k} \right)^2 \right] \\ & - 2b \left[SF\sigma_1^2 - SF\sigma_1 \left(\frac{\sigma_2(1 - e^{-k\tau})}{k} \right) \rho_{12} - F\sigma_3 \left(\frac{\sigma_2(1 - e^{-k\tau})}{k} \right) \rho_{23} + F\sigma_1\sigma_3\rho_{13} \right] \end{aligned} \quad 4.23$$

The differences between expressions 4.17 and 4.23 are in the terms that represents the instantaneous variance of cash price changes and the covariance between cash and futures price changes. The variance of cash price changes now reflects the interaction between the spot price and its difference with the local cash price. The covariance term now contains portions that reflect the covariation of B with the other stochastic factors in the model. This results in an expression for the optimal hedge ratio, analogous to equation 4.18, of

$$b_{AGS} = \frac{-\lambda_U^{-1} \mu_F + \left[SF\sigma_1^2 - SF\sigma_1 \left(\frac{\sigma_2(1-e^{-k\tau})}{k} \right) \rho_{12} - F\sigma_3 \left(\frac{\sigma_2(1-e^{-k\tau})}{k} \right) \rho_{23} + F\sigma_1\sigma_3\rho_{13} \right]}{F^2 \left[\sigma_1^2 - 2\sigma_1\sigma_2\rho \left(\frac{(1-e^{-k\tau})}{k} \right)_{12} + \sigma_2^2 \left(\frac{(1-e^{-k\tau})}{k} \right)^2 \right]}$$

4.24

This is a more general optimal hedge ratio that could be used by a hedger who does not plan to make delivery at the delivery location specified by the futures contract, or who is implementing a cross hedge.

Schwartz (1997) presents a one-factor model for pricing commodity contingent claims, hereafter referred to as the S97 model. Rather than arguing that a risk-free portfolio of a derivative and the underlying commodity can be formed, however, this model is developed by attaching a market price of (spot price) risk to the derivative. The S97 model does not therefore follow in the spirit of Kaldor, Working and Brennan's theory of storage as the GS model did, but instead follows Keynes (1930) and Hicks (1939) in emphasizing the role of risk and return in determining the value of contingent claims. In the S97 model, the spot price is assumed to follow the process

$$dS = k(\mu - \ln S)Sdt + \sigma_1 Sdz_1. \quad 4.25$$

where as before σ_1^2 is the instantaneous variance of changes in the natural logarithm of the spot price, and the log of the spot price reverts to level μ at rate k . The price of a futures contract must satisfy, as discussed by Schwartz, the partial differential equation

$$\frac{1}{2}\sigma_1^2 S^2 F_{SS} + k(\mu - \lambda - \ln S)SF_S - F\tau \quad 4.26$$

where λ is the market price of risk. Schwartz gives the solution as

$$F(S, \tau) = \exp \left[e^{-k\tau} \ln S + (1 - e^{-k\tau}) \left(\mu - \frac{\sigma_1^2}{2k} - \lambda \right) + \frac{\sigma_1^2}{4k} (1 - e^{-2k\tau}) \right], \quad 4.27$$

Following the discussion of hedging using the GS model, when $C = S$, we have the following process for the short hedger's portfolio under the S97 model

$$dP = \mu_P dt + \sigma_P dz \quad 4.28$$

where the drift is

$$\mu_P = k(\mu - \ln S)S - bFe^{-k\tau}k\lambda \quad 4.29$$

and instantaneous variance is

$$\sigma_P^2 = S^2 \sigma_1^2 + b^2 F^2 e^{-2k\tau} \sigma_1^2 - 2bSFe^{-k\tau} \sigma_1^2. \quad 4.30$$

The optimal hedge ratio for the short hedger when $L = S$ is then

$$b_{S97} = \frac{-\lambda_U^{-1} k\lambda + S\sigma_1^2}{Fe^{-k\tau} \sigma_1^2}. \quad 4.31$$

Note that if we ignore the speculative component, the variance-minimizing hedge ratio is $(S/F)\exp(k\tau)$. Using equation 4.27, it is easy to see that this is identical to F_S . This demonstrates that the adaptation of contingent claims models for mean-variance hedging that we outline here can be considered a generalization of delta hedging. Augmenting the S97 for the case where $L \neq S$, again using equation 4.20 and specifying

$$dB = \gamma(\beta - B)dt + \sigma_2 dz_2, \quad 4.32$$

similar to before we find the diffusion followed by the hedger's portfolio is

$$dP_A = \mu_{AP} dt + \sigma_{AP} dz \quad 4.33$$

with drift

$$\mu_{AP} = \gamma(\beta - B) + k(\mu - \ln S)S - bFe^{-k\tau}k\lambda \quad 4.34$$

and instantaneous variance

$$\sigma_{AP}^2 = [S^2\sigma_1^2 + \sigma_2^2 + 2S\sigma_1\sigma_2\rho_{12}] + b^2[F^2e^{-2k\tau}\sigma_1^2] - 2h[SF e^{-k\tau}\sigma_1^2 + Fe^{-k\tau}\sigma_1\sigma_2\rho_{12}] \quad 4.35$$

The optimal hedge ratio for the short hedger when $L \neq S$ is then given by

$$b_{AS97} = \frac{-\lambda_U^{-1}k\lambda + S\sigma_1^2 + \sigma_1\sigma_2\rho_{12}}{Fe^{-k\tau}\sigma_1^2}. \quad 4.36$$

DATA, PARAMETER ESTIMATION, AND PARAMETER INFERENCE

The data we use are week ending observations of the New York Mercantile Exchange (NYMEX) crude oil futures contracts, options on those futures, and the associated spot price. The futures and spot price data are observed over the period January 6, 1984 through June 21, 2002. We use option prices observed January 3, 1992 through June 21, 2002. Option prices were available before 1992, but trading volumes were not sufficient for the purposes outlined below. All data were provided by Commodity Research Bureau. We divide the data into three periods. The first time period, January 6, 1984 through December 27, 1991 (417 observations), is used strictly for parameter estimation. The second time period, January 3, 1992 through December 27, 1996 (261 observations), is used for both parameter estimation and the evaluation of

in-sample hedging effectiveness.¹¹ Out-of-sample hedging effectiveness is evaluated over the final time period, January 3, 1997 through June 21, 2002 (286 observations).

There is one NYMEX crude oil futures delivery per month. The price data for individual futures contracts were used to construct a rolling nearby futures series (*NEAR*) that is used in the parameter estimation and evaluation of hedging effectiveness. Where price changes were required, as in the unit root testing and VEC model estimation, care was taken to take changes of the individual futures series before selecting those changes that were nearby. That is to say, we use nearby futures changes (*NEARD*) rather than a changes in the nearby futures series (*DNEAR*). The latter series would result in roughly one out of every four observations being the composition of a change in a futures price and the spread between the expiring and new nearby futures prices (due to monthly contract expiration and the weekly observation frequency). Such a series has no natural interpretation in the context of hedging, and an uncertain (at best) interpretation in the context of time series econometrics. The *NEARD* series, however, contains no observations that are corrupted by futures spreads and is consistent with the futures price changes that an actual trader would realize. The differenced spot price series (*DS*) contains the usual first differences of the spot prices (*S*).

Following Gibson and Schwartz (1990), we employ the annualized one-month forward convenience yield when estimating the stochastic processes underlying the GS

¹¹ In-sample hedging effectiveness is not evaluated over the entire in-sample estimation period because option trading volume was insufficient to carry out the inference of the term structure of volatility in the S97 model.

model. This is estimated using the price F^1 of a nearby futures contract and the price F^2 of the subsequent contract expiring using the following relation

$$\delta = r^1 - 12 \ln \left(\frac{F^1}{F^2} \right) \quad 4.37$$

where r^1 is the one-month forward riskless interest rate.

We first discuss the in-sample time series analysis. Augmented Dickey-Fuller (ADF) Tests for unit roots were carried out on all series over the in-sample estimation period (January 6, 1984 through December 27, 1996), with results presented in the first four rows of table 4.1. Test results suggest non-stationary behavior, and differenced spot and nearby futures changes series are thus used for the remainder of the time-series estimation. We test for the presence of cointegration between S and $NEAR$ using the Engle-Granger (1987) methodology.¹² Regressing S on $NEAR$ and a constant results in the following potential cointegrating relation

$$ECT = S + 0.014 - 1.001NEAR. \quad 4.38$$

An ADF test statistic on the recovered ECT series, presented in the last row of table 4.1, strongly rejects the null hypothesis of a unit root, and we conclude that S and $NEAR$ are indeed cointegrated.

¹² Unfortunately, available implementations of Johansen's (1988) cointegration methodology perform data differencing automatically when forming the vector auto-regression. In the present context, given the series $NEAR$, an implementation of the Johansen methodology would then generate and subsequently employ the unacceptable differenced nearby series $DNEAR$ described above. Hypothesis testing on the coefficients of the cointegrating vector within the Engle-Granger framework can be misleading (Stock 1987), however we carry out no such testing. The Engle-Granger methodology does provide a consistent estimate of a single cointegrating vector, however, which is all that we require here.

Preliminary univariate analysis of the *DS* and *NEARD* series suggested the presence of GARCH effects as expected. Bollerslev's GARCH(1,1) process was then fitted to each series under the assumption of normality, with the results found in table 4.2. Consistent with Baillie and Myers, no autoregressive terms in the mean equations were necessary to render the standardized residuals free of autocorrelation, as evidenced by the reported Ljung-Box tests on the standardized residuals for up to 12th-order autocorrelation. The sample skewness and kurtosis of the standardized residuals from each model suggest no significant deviation from normality. Asymptotic standard errors for the conditional variance equation parameter estimates confirm the presence of GARCH behavior in the series, and the Ljung-Box test on the squared standardized residuals indicates that the GARCH(1,1) specification adequately represents this behavior.

Based on the results of the univariate time series analyses, the multivariate VEC-GARCH(1,1) model given by equations 4.5 through 4.7 was fitted to the *DS* and *NEARD* series under the assumption of normality. The mean equations for each variable include the *ECT* recovered using equation 4.38. Schwarz (1978) information criterion was employed in the specification of the mean equations otherwise, and it was determined that neither constants nor autoregressive terms were desirable. Results are presented in table 4.3. Residual diagnostics suggest no serious misspecification. All parameter estimates are significant at the 1% level. The speed of adjustment coefficients on the *ECT* suggest that deviations from the long-run equilibrium are subject to rapid correction, as expected given the frequency of futures deliveries used to construct the

NEARD series. The parameters estimates associated with the conditional variance dynamics (A_{ii}, B_{ii}, W_i ; $i = 1, 3$) are similar to those obtained in the univariate estimation.

The parameter estimates associated with the conditional covariance dynamics (A_{22}, B_{22}, W_2) indicate substantial interaction between the two series.

The GS model parameters were estimated using an iterated seemingly unrelated regressions (SUR) procedure on the linear discrete approximations to equations 4.8 and 4.9. The resulting annualized parameter estimates are $\mu = -0.017$, $\alpha = 0.177$, $k = 9.183$, $\sigma_1 = 0.349$, $\sigma_2 = 1.157$, and $\rho_{12} = 0.431$. The large estimate of k suggests a high degree of mean-reversion in the convenience yield, and the large estimate of σ_2 suggests that it is highly volatile as well. We refer to this method of parameter estimation as estimating the stochastic differential equations (SDEs).

In order to implement the optimal hedging scheme outlined in section III, an estimate of the market price of convenience yield risk in the GS model is also needed. To accomplish this task, we follow Gibson and Schwartz (1990) by finding the least-squares fit of the futures pricing formula in the GS model to the market data. Specifically, for each available week-ending futures price observation for each delivery in the data set over the in-sample period, we collect the 5-tuple (F, S, δ, τ, r) . We then use all such observations to find the value of λ that minimizes the sum of squared pricing errors implied by equation 4.12, using the estimates of parameters other than λ found by estimating the SDEs. The value of λ that we find is -0.132 . As discussed in Gibson and Schwartz (1990), finding a negative price of convenience yield risk is

consistent with the fact that the partial derivative of the futures price with respect to the convenience yield is negative.

In addition to estimating the SDEs, it is also possible to directly estimate the parameters of the term structure of volatility (TSV) in the GS model, using market data observed during the recent past. This provides a means by which the restrictive assumption of a constant TSV can be somewhat relaxed. The TSV for the GS model is given by

$$\sigma_F(\tau; \sigma_1, \sigma_2, \rho_{12}, k) = \sqrt{\sigma_1^2 - 2\sigma_1\sigma_2\rho_{12}\left(\frac{1-e^{-k\tau}}{k}\right) + \sigma_2^2\left(\frac{1-e^{-k\tau}}{k}\right)^2} \quad 4.39$$

Computing the annualized sample standard deviations of observed futures log price changes for the most recent 2 months of daily observations for the n th nearby futures series provides us with a pair $(\hat{\sigma}_F, \tau)$ where τ is the average length of time until expiration. Collecting these pairs for the 12 nearest nearby futures price series provides 12 observations with which we find the values of $\sigma_1, \sigma_2, \rho_{12}$ and k that result in the best fit, in the least squares sense, of equation 4.39 to the market data. This exercise can be carried out at any point in time to arrive at a TSV that reflects more recent market activity, rather than a very long run average TSV found by estimating the SDEs. The estimated TSV might be thought of as the generalization of what is commonly referred to as “historical volatility.” Rather than estimating the annualized volatility of only the spot price using a moving window of observations, however, the entire TSV is estimated. This provides a second means that a hedger might use arrive at the GS parameters needed to calculate his optimal hedge ratio. As an example, figure 4.1

presents the GS term structure of volatility found by estimating the SDEs, and the TSV found by direct estimation on June 21, 1996 (a date chosen to illustrate an example of a high level of volatility in nearby futures). In both cases, the TSV is a decreasing function of time until maturity, as predicted by the Samuelson hypothesis. The functional form for the TSV in the GS model does not require this, however; gentle increases at longer times until maturity are permitted and are observed over some intervals in the data set.

In addition to the two parameter estimation methods discussed above, it is also theoretically possible to infer the TSV from observed futures option prices if a closed-form solution for those prices is available for a given model. In the case of the GS model, the value C at time t , of a European call option with strike price X , expiring at time T_1 , on a futures contract with price F , expiring at time T , is given in Hilliard and Reis (1998) as

$$C(F, X, t, T_1, T) = e^{-r(T_1-t)} [FN(d_1) - XN(d_1 - v)] \quad 4.40$$

where

$$d_1 = \frac{\ln(F/X) + 0.5v^2}{v} \quad 4.41$$

and

$$\begin{aligned} v^2(t, T_1, T) = & \sigma_1^2 (T_1 - t) - \frac{2\sigma_1\sigma_2\rho_{12}}{k} \left[(T_1 - T) - \frac{e^{-k(T-T_1)} - e^{-k(T-t)}}{k} \right] \\ & + \frac{\sigma_2^2}{k^2} \left[(T_1 - t) - \frac{2}{k} (e^{-k(T-T_1)} - e^{-k(T-t)}) + \frac{1}{2k} (e^{-2k(T-T_1)} - e^{-2k(T-t)}) \right]. \end{aligned} \quad 4.42$$

$N(d_1)$ represents the standard normal distribution function evaluated at d_1 . To infer the TSV on a given date, the price for one approximately at-the-money option on each futures contract was collected (where available). All such available option prices and the corresponding values of F , X , r , T , and T_1 , were then used in attempts to find the values of $\sigma_1, \sigma_2, \rho_{12}$ and k that provided the best least-squares fit of (the highly non-linear) equation 4.40. Just as the direct estimation of the TSV can be thought of as a generalization of “historical volatility”, the option-implied TSV can be thought of as a generalization of “implied volatility”. Unfortunately, in many cases as few as 5 observations were available for this task, and the inferred parameter values were often unreasonable. Given that this task could not be performed reliably with the available data, we do not use option-implied term structures of volatility for hedging in the context of the GS model.

We now turn to the estimation of the parameters of the S97 model. The linear discrete approximation of equation 4.25 was estimated over the in-sample estimation period using ordinary least squares, resulting in the following annualized parameter estimates: $\mu = 3.038$, $\alpha = 2.993$, $k = 1.334$, and $\sigma_1 = 0.347$. The market price of risk in the S97 model was estimated using a procedure analogous to that used to estimate the market price of convenience yield risk in the GS model. The resulting in-sample estimate of the market price of risk λ is 0.025. In addition to estimating the SDE of the S97 model, it is again possible to directly estimate the TSV. The TSV for the S97 model is given by

$$\sigma_F(\tau; \sigma_1, k) = e^{-k\tau} \sigma_1. \quad 4.43$$

Again pairs $(\hat{\sigma}_F, \tau)$ were collected for the 12 nearest nearby futures series, and the natural logarithm of $\hat{\sigma}_F$ was regressed on τ to arrive at least squares estimates for k and σ_1 . In the case of the S97 model, we find that it is possible to reliably infer the TSV using observed futures option prices. The solution for European options on futures in the S97 model is given in Clewlow and Strickland (1999). The solution is equations 4.40 and 4.41 again, but equation 4.42 is replaced with

$$v^2(t, T_1, T) = \frac{\sigma_1^2}{2k} \left[e^{-2k(T-T_1)} - e^{-2k(T-t)} \right]. \quad 4.44$$

The term structures of volatility estimated/inferred using the three methods outlined above for the S97 model on June 21, 1996 are presented in figure 4.2. Note first that in all cases the TSV is a strictly decreasing function of time until maturity as dictated by its exponential decay functional form. The directly estimated TSV indicates a higher level of volatility at all times until maturity than the option-implied TSV. As it happened, the option-implied TSV indicated much higher levels of volatility one or two months earlier. This highlights the lagged effect that an increase in the general level of volatility will have on the TSV that is directly estimated using a moving window of historical data. The option-implied TSV, on the other hand, is calculated using data observed on a single day and can therefore adjust instantly to changes in market conditions.

Careful examination of the dynamics of the implied TSV, however, reveals a more subtle problem. We found the S97 option-implied TSV displayed a teetering behavior – an increase in implied spot price volatility was generally accompanied by a

decrease in the implied volatility of futures far from maturity and vice versa. Evidence of this is presented in figure 4.3. Over a six week period, the implied spot price volatility increased roughly 8%, while the implied volatility of futures one year from expiration decreased about 4%. This phenomenon seems difficult to justify economically, and more likely result from the assumption of a constant TSV. In actual practice, option traders anticipate mean reversion in volatility levels - an increase in spot price volatility is likely to die out as time passes. As discussed in Hull and White (1987), the prices of options in a stochastic volatility environment should be a function of the expected levels of volatility over the life of the option. A short-term increase in spot price volatility has a large impact on the average level of volatility over the life of option that is nearing expiration, but a relatively small impact on the average level of volatility expected over the life of an option far from expiration. A significant increase in the premiums for options on nearby futures may therefore be accompanied by only a modest increase in the prices of options on distant futures. A significant increase in nearby option prices necessarily results in an increase in the value of σ_1 in the fitted TSV, but the rate of decay of volatility k must also increase if the distant option prices have not risen by much.

HEDGING EFFECTIVENESS

We consider the problem of a hypothetical crude oil trader with mean-variance utility that wishes to take an optimal position in crude oil futures using equation 4.4. We assume that the cash position is 100,000 barrels, and that this position is hedged using

the nearby futures contract. We further assume that the hedger's cash position corresponds to the specifications of the futures contract (i.e. $L = S$). Optimal hedge ratios in the time series hedging scheme are formed in each period by using the appropriate elements of the conditional variance-covariance matrix H_t . When employing partial equilibrium hedging, hedge ratios are formed using either equation 4.18 or equation 4.31 after any appropriate parameter estimation or inference. Two methods of parameter estimation are devised above for the GS model: 1) simply estimating the SDEs and 2) directly estimating the TSV each time a new hedge ratio is formed. These two methods of parameter estimation are also available when using the S97 model, and we additionally are able to infer the TSV from futures option prices. We thus have five competing partial equilibrium hedging schemes.

The time paths of hedge ratios generated by the VEC-GARCH and GS models during the last 18 months of the in-sample period are presented in figure 4.4. As expected, the hedge ratios generated by the GS model with SDE parameters estimates are fairly stable relative to those generated by the VEC-GARCH and GS model employing a freshly estimated TSV each period. Nonetheless, the paths of the VEC-GARCH and GS with SDE parameter estimates are similar – steady in late 1995, then dipping in the spring and summer of 1996 and then increasing in late 1996. Over the portion of the in-sample period for which we evaluate hedging effectiveness, the correlation coefficient between these two models' hedge ratios is 0.53, while the correlation between the hedge ratios from the VEC-GARCH and the GS model with an estimated TSV is -0.10 . The time paths of the optimal hedge ratios generated by the

three S97 models over the same time period are presented in figure 4.5. All three consistently follow a saw tooth pattern due to the functional form of the TSV in the S97 model. Ignoring the speculative component of equation 4.31, and assuming the ratio of the spot price to the futures price is approximately one, the optimal hedge ratio is then approximately $\exp(k\tau)$. This is greater than one before futures expiration, and decays to one at the time of expiration. As one might expect after examining figures 4.4 and 4.5, the S97 hedge ratios are highly correlated with one another, but not with the GS or VEC-GARCH hedge ratios.

To evaluate in-sample hedging effectiveness, the realized levels of certainty equivalent income (CEI), based on the realized price changes and conditional variances and covariances from the VEC-GARCH model, were evaluated for each week over the period January 3, 1992 through December 27, 1996. The average level of CEI was then calculated for each of the six hedging models. Table 4.4 presents these averages, as well as that realized for the unhedged cash position. CEI increases are large in all cases, demonstrating the excellent hedging performance of NYMEX crude oil contract in the present context. The VEC-GARCH model delivers the greatest utility increase, at 85.16%. Among the partial equilibrium models, there is no clear-cut pattern. Neither the GS nor S97 models' performance dominates the other. Also, neither of the two available methods of parameter estimation is clearly superior. Hedging using the GS model with estimated SDEs results in hedging performance that is very similar to hedging using the S97 model with estimated term structures of volatility. The S97 model with option-implied terms structures of volatility provides the second worst

hedging performance, despite the attempt to glean insight into the future volatility conditions expected by option traders.

Previous optimal hedging literature considers not only in-sample hedging effectiveness, but stresses the need to evaluate out-of-sample hedging effectiveness as well. This provides a fair test of how an optimal hedging scheme is likely to perform in real-world conditions. To evaluate out-of-sample hedging effectiveness, we re-estimated each model each period using all available data at that point in time for each of the models, and then used each to make one period ahead forecasts of the components of the hedger's optimal hedge ratio. The resulting CEI in each period was assessed using the ensuing actual price changes in the following week and the conditional variances and covariances recovered from a final VEC-GARCH model estimated using the entire data set. Again the CEIs from each period were averaged for each hedging model and for the unhedged case. Results are presented in table 4.5. These results are very similar to those found in the in-sample period. The VEC-GARCH model results in the largest CEI increase. Again the S97 model with estimated TSV and the GS model with estimated SDEs deliver similar performance, roughly tying for second place. The remaining three models again share the dishonor of being the three worst performing.

To determine if the superior hedging effectiveness of the VEC-GARCH model was attributable to superior futures price forecasting (associated with the speculative component of the hedge ratio) or the superior variance and covariance forecasting, the out-of-sample forecasts of nearby futures one-week price changes were evaluated. All models delivered very similar root mean squared forecast errors (RMSEs), however the

VEC-GARCH model provided the worst forecasts. The RMSE of the VEC-GARCH forecasts was \$1.161 per barrel, while the partial equilibrium models' RMSEs were tightly distributed around an average of \$1.154 per barrel. It therefore appears that the superior hedging performance of the VEC-GARCH model is due entirely to superior modeling of conditional variance and covariance dynamics.

Overall, the VEC-GARCH hedging model, which allows time-varying variances and covariance, provides the best hedging performance, despite producing the most variable hedge ratios (as measured by sample standard deviation). The partial equilibrium models' hedge ratios are less variable, but perform worse. This suggests that the hedge ratios generated by the partial equilibrium models are not sufficiently reflecting changes in volatility conditions. The cause of the inferior performances of the partial equilibrium hedging models thus appears to be the unrealistic assumption of a constant TSV. Attempts to compensate for this shortcoming by frequently estimating or inferring the TSV do not result in consistently improved hedging effectiveness, and in no case is the performance of the VEC-GARCH model matched. Estimating the TSV suffers from the problem of employing a moving window of historical data, and any change in volatility conditions is reflected with somewhat of a lag. Inferring the TSV from futures options prices (only practical for the S97 model) is still done in a constant TSV context, and suffers from the teetering effect described earlier. All methods of updating the parameters of the term structures of volatility in the partial equilibrium models also come at the expense of a significant increase in computational complexity.

In a sense, the hedging problem formulated here was the easiest possible for the partial equilibrium models. The assumption was made that the hedger's cash position corresponded with the futures contract specifications (i.e. $L = S$). We thus employed the optimal hedge ratios in equations 4.18 and 4.31 rather than those from the augmented models in equations 4.24 and 4.36. For many hedgers this will not be the case, and the use of the augmented models would be necessary. This would likely result in hedging performance that fell further short of that of the VEC-GARCH model, for the following reason. The use of an augmented partial equilibrium models would add another layer of constant variance-covariance assumptions – likely exacerbating the problem that led to the poor performance when $L = S$. On the other hand, the case where $L \neq S$ presents no special problem for the time series model, as one would simply employ the appropriate local cash price series rather than the spot price series, and proceed as usual with a model that fully incorporates conditional variance and covariance dynamics.

CONCLUSIONS

This chapter compares the performances of time series and partial equilibrium based optimal hedging models for trader that is long in a cash commodity market, and maximizes mean-variance utility using futures contracts. We find that the time series approach delivers superior hedging performance to that of each of the other models considered. This appears to be due to the partial equilibrium models' unpalatable assumption of a constant volatility term structure. The constant volatility term structure

framework hampers even the seemingly promising technique of inferring option market participants' expectations regarding future volatility conditions.

This research has considered only a single type of derivative, however. These results suggest that the attractiveness of employing a simple partial equilibrium model (i.e. one that does not incorporate stochastic volatility) when hedging a commodity market cash position using futures contracts (or *vice versa*) is questionable. Few would doubt the usefulness of partial equilibrium models in hedging a position in a derivative with a non-linear payoff function (e.g. an option), however. The conclusion then is that different types of hedging models are suited to different tasks, and the best approach in still other situations (e.g. hedging with swaps) is uncertain. Furthermore, this research has considered only a single hedging objective. When commodity producers or consumers purchase options they generally think of them as being similar to insurance contracts. This suggests that they may be maximizing utility of a form other than that employed here (and in much of the optimal hedging literature). These issues illuminate the necessity of further research.

CHAPTER V

CONCLUSIONS

This dissertation examines issues regarding the workings and uses of futures markets. Given the important roll that futures trading plays in the economy, a sound understanding of their workings and uses is critical. This dissertation makes contributions towards such understanding in three areas.

Chapter II evaluates the relative performances of various estimators of bid-ask spreads in futures markets using commonly available transaction data. Results indicate a wide divergence in the performances of the competing estimators. Absolute price change estimators generally perform better than serial covariance estimators. Encompassing test indicate, however, that composite estimators may be useful. This chapter also examines the effect of automating trading on spreads in commodity futures markets. Results indicate that spreads generally widened after trading was automated on the markets considered, and the tendency for spreads to widen during periods of high volatility increased. These results are in contrast to those found in higher volume financial futures markets, demonstrating the necessity of considering lower-volume commodity futures markets separately. The results also call into question the advisability of automating trading on these exchanges.

In Chapter III investigates various unresolved issues regarding futures markets, using formal methods appropriate for inferring causal relationships from observational data when some relevant quantities are hidden. I find no evidence supporting the

generalized version of Keynes's theory of normal backwardation. I find no evidence supporting theories that predict that the level of activity of speculators or uninformed traders affects the level of price volatility, either positively or negatively. My evidence strongly supports the mixture of distribution hypothesis (MDH) that trading volume and price volatility have one or more latent common causes, resulting in their positive correlation. There are abundant opportunities for the further application of causal inference methods to empirical research into derivatives markets. Other open questions need to be addressed, some of which are: is the level of futures trading activity a cause of price volatility in the underlying cash market? What are the causes and/or effects of changes in the shape of the forward curve? What are the causes of basis movements? Does the size of the margin deposit required to trade futures impact any of the quantities that we have considered? What are the causal relationships that exist across related markets (e.g. the soy complex or the crude oil complex)? What are the causal relationships across time that exist among the variables that we considered? This chapter represents only the beginning of the exploration of derivative markets issues using these exciting new techniques.

In Chapter IV compares the performances of the partial equilibrium and statistical approaches to hedging. Different types of hedgers have traditionally used each of two approaches: derivatives dealers and market makers have typically used the former approach to hedge their portfolios, while commodity producers and consumers more commonly use the latter. This research provides the first known comparison of the out-of-sample hedging performance of the two approaches. Results indicate that for a

simple derivative with a linear payoff function (a futures contract), the statistical models significantly outperform the partial equilibrium models considered here. This chapter considers only this single type of derivative. Few would doubt the usefulness of partial equilibrium models in hedging a position in a derivative with a non-linear payoff function (e.g. an option), however. The conclusion then is that different types of hedging models are suited to different tasks, and the best approach in still other situations is uncertain. Furthermore, this research has considered only a single hedging objective. When commodity producers or consumers purchase options they generally think of them as being similar to insurance contracts. This suggests that they may be maximizing utility of a form other than that employed here (and in much of the optimal hedging literature). These issues await further investigation.

BIBLIOGRAPHY

- Akleman, D. G., Bessler, D. A., & Burton, D. M. (1999). Modeling Corn Exports and Exchange Rates with Directed Graphs and Statistical Loss Functions. In C. G. and G. F. Cooper (Eds.), *Computation, Causation, and Discovery*, Cambridge, MA: MIT Press.
- ap Gwilym, O., & Thomas, S. (in press). An Empirical Investigation of Quoted and Implied Bid-Ask Spreads on Futures Contracts. *Journal of International Financial Markets, Institutions, and Money*.
- Ashley, R., Granger, C. W. J., & Schmalensee, R. (1980). Advertising and Aggregate Consumption: An Analysis of Causality. *Econometrica*, 48, 1149-1167.
- Bae, K., Chan, K., & Cheung, Y. (1998). The Profitability of Index Futures Arbitrage: Evidence from Bid-Ask Quotes. *Journal of Futures Markets*, 18, 743-763.
- Ballie, R. T., & Myers, R. J. (1991). Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge. *Journal of Applied Econometrics*, 6, 109-124.
- Bessimbinder, H. (1992). Systematic Risk, Hedging Pressure, and Risk Premiums in Futures Markets. *Review of Financial Studies*, 5, 637-667.
- Bhattacharya, M. (1983). Transactions Data Tests of Efficiency of the Chicago Board of Options Exchange. *Journal of Financial Economics*, 12, 161-186.
- Black, F. (1976). The Pricing of Commodity Contracts. *Journal of Financial Economics*, 3, 167-179.

- Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81, 637-654.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307-327.
- Bollerslev, T., Engle, R. F., & Wooldridge, J. M. (1988). A Capital Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy*, 96, 116-131.
- Brandt, J. A., & Bessler D. A. (1983). Price Forecasts and Evaluation: An Application in Agriculture. *Journal of Forecasting*, 2, 237-248.
- Brennan, M. J. (1958). An Approach to the Valuation of Uncertain Income Streams. *American Economic Review*, 48, 50-72.
- Brennan, M. J. (1991). The Price of Convenience and the Valuation of Commodity Contingent Claims. In D. Lund and B. Oskendal (Eds.), *Stochastic Models and Option Values*. Amsterdam, Netherlands: North-Holland.
- Brennan, M. J., & Schwartz, E. S. (1985). Evaluating Natural Resource Investments. *The Journal of Business*, 58, 135-157.
- Carter, C. A., Rausser G. C., & Schmitz, A. (1983). Efficient Asset Portfolios and the Theory of Normal Backwardation. *Journal of Political Economy*, 91, 319-331.
- Cecchetti, S. G., Crumby, R. E., & Figlewski, S. (1988). Estimation of the Optimal Futures Hedge. *Review of Economics and Statistics*, 70, 623-630.
- Chang, E. C. (1985). Returns to Speculators and the Theory of Normal Backwardation. *Journal of Finance*, 40, 193-208.

- Chang, E. C., Chou, R. Y., & Nelling, E. F. (2000). Market Volatility and the Demand for Hedging in Stock Index Futures. *Journal of Futures Markets*, 20, 105-125.
- Chu, Q. C., Ding, D. K., & Pyun, C. S. (1996). Bid-Ask and Spreads in the Foreign Exchange Market. *Review of Quantitative Finance and Accounting*, 6, 19-37.
- Clark, P. K. (1973). A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica*, 41, 135-155.
- Clelow, L., & Strickland, C. (1999). Valuing Energy Options in a One-Factor Model Fitted to Forward Prices. Working paper, University of Technology.
- Cootner, P. H. (1960). Returns to Speculators: Telser vs. Keynes. *Journal of Political Economy*, 68, 396-404.
- Copeland, T. E., & Galai, D. (1983). Information Effects on the Bid-Ask Spread. *Journal of Finance*, 38, 1457-1469.
- Daigler, R. T., & Wiley, M. K. (1999). The Impact of Trader Type on the Futures Volatility-Volume Relation. *Journal of Finance*, 54, 2297-2316.
- Danthine, J. (1978). Information, Futures Prices, and Stabilizing Speculation. *Journal of Economic Theory*, 17, 79-98.
- DeJong, D. N., Nankervis, J. C., Savin, N. E., & Whiteman, C. H. (1992). Integration Versus Trend Stationary in Time Series. *Econometrica*, 60, 423-433.
- DeLong, J. B., Shleifer, A., Summers, L., & Waldman, R. J. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy*, 98, 703-738.
- Demiralp, S., & Hoover, K. D. (2003). Searching for the Causal Structure of a Vector Autoregression. Working Paper 03-03, University of California, Davis.

- Demsetz, H. (1968). The Cost of Transacting. *Quarterly Journal of Economics*, 82, 33-53.
- Ding, D. K. (1999). The Determinants of Bid-Ask Spreads in the Foreign Exchange Futures Markets: A Microstructure Analysis. *Journal of Futures Markets*, 19, 307-324.
- Dusak, K. (1973). Futures Trading and Investor Returns: An Investigation of Commodity Market Risk Premiums. *Journal of Political Economy*, 81, 1387-1406.
- Ederington, L., & Lee, J. H. (2002). Who Trades Futures and How: Evidence from the Heating Oil Futures Market. *Journal of Business*, 75, 353-373.
- Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica*, 50, 987-1008.
- Engle, R. F., & Granger, C. W. J. (1987). Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica*, 55, 251-276.
- Epps, T. W., & Epps, M. L. (1976). The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture of Distributions Hypothesis. *Econometrica*, 44, 305-321.
- Foster, F. D., & Viswanathan, S. (1994). Strategic Trading with Asymmetrically Informed Traders and Long-Lived Information. *Journal of Financial and Quantitative Analysis*, 29, 499-518.

- Frino, A., McInish, T. H., & Toner, M. (1998). The Liquidity of Automated Exchanges: New Evidence from German Bund Futures. *Journal of International Financial Markets, Institutions, and Money*, 8, 225-241.
- Fuller, W. A. (1976). *Introduction to Statistical Time Series*. New York, NY: Wiley.
- Gagnon, L., Lypny, G. J., & McCurdy, T. H. (1998). Hedging Foreign Currency Portfolios. *Journal of Empirical Finance*, 5, 197-220.
- Gibson, R., & Schwartz, E. S. (1990). Stochastic Convenience Yield and the Pricing of Oil Contingent Claims. *Journal of Finance*, 45, 959-976.
- Gibson, R., & Schwartz, E. S. (1991). Valuation of Long Term Oil-Linked Assets. In D. Lund and B. Oskendal (Eds.), *Stochastic Models and Option Values*. Amsterdam, Netherlands: North-Holland.
- Gibson, R., & Schwartz, E. S. (1993). The Pricing of Crude Oil Futures Options Contracts. *Advances in Futures and Options Research*, 6, 291-311.
- Glosten, L. R. & Milgrom, P. R. (1985). Bid , Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders. *Journal of Financial Economics*, 14, 71-100.
- Granger, C. W. J., & Newbold, P. (1973). Some Comments on the Evaluation of Economic Forecasts. *Applied Economics*, 5, 35-47.
- Haigh, M. S., & Bessler, D. A. (in press). Causality and Price Discovery: An Application of Directed Acyclic Graphs. *Journal of Business*.
- Haigh, M. S., & Holt, M. T. (2000). Hedging Multiple Price Uncertainty in International Grain Trade. *American Journal of Agricultural Economics*, 65, 603-605.

- Hartzmark, M. L. (1987). Returns to Individual Traders of Futures: Aggregate Results. *Journal of Political Economy*, 95, 1292-1306.
- Harvey, D. I., Leybourne, S. J., & Newbold, P. (1998). Test for Forecast Encompassing. *Journal of Business & Economic Statistics*, 16, 254-259.
- Hey, J. D. (1979). *Uncertainty in Microeconomics*. New York, NY: New York University Press.
- Hicks, J. R. (1939). *Value and Capital*. Oxford, England: Oxford University Press.
- Hilliard, J. E., & Reis, J. (1998). Valuation of Commodity Futures and Options Under Stochastic Convenience Yields, Interest Rates, and Jump Diffusions in the Spot. *Journal of Financial and Quantitative Analysis*, 33, 61-86.
- Hirschleifer, D. (1988). Residual Risk, Trading Costs, and Commodity Futures Risk Premia. *Review of Financial Studies*, 1, 173-193.
- Hirschleifer, D. (1990). Hedging Pressure and Futures Price Movements in a General Equilibrium Model. *Econometrica*, 58, 411-428.
- Houthakker, H. S., (1957). Can Speculators Forecast Prices?" *Review of Economics and Statistics*, 39, 143-151.
- Hull, J. C., & White, A. (1987). The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance*, 42, 281-300.
- Johansen, S. (1988). Statistical Analysis of Cointegrating Vectors. *Journal of Economics Dynamics and Control*, 52, 169-210.
- Johnson, L. L. (1960). The Theory of Hedging and Speculation in Commodity Futures. *Review of Economic Studies*, 27, 139-151.

- Kahl, K. H. (1983). Determination of the Recommended Hedging Ratio. *American Journal of Agricultural Economics*, 65, 603-605.
- Kaldor, N. (1939). Speculation and Economics Stability. *Review of Economic Studies*, 7, 1-27.
- Karpoff, J. M. (1987). The Relation Between Price Changes and Trading Volume: A Survey. *Journal of Financial and Quantitative Analysis*, 22, 109-126.
- Keynes, J. M. (1930). *A Treatise on Money*. London, England: Macmillan.
- Kroner, K. F., & Sultan, J. (1993). Time Varying Distributions and Dynamic Hedging with Foreign Currency Futures. *Journal of Financial and Quantitative Analysis*, 28, 535-551.
- Laux, P.A., & Senchack, A. J. (1992). Bid-Ask Spreads in Financial Futures. *Journal of Futures Markets*, 12, 621-634.
- Lee, C., & Ready, M. (1991). Inferring Trade Direction from Intraday Data. *Journal of Finance*, 46, 733-746.
- Locke P. R., & Venkatesh, P. C. (1997). Futures Market Transaction Costs. *Journal of Futures Markets*, 17, 229-245.
- Ma, C. K., Peterson, R. L., & Sears, R. S. (1992). Trading Noise, Adverse Selection, and Intraday Bid-Ask Spreads in Futures Markets. *The Journal of Futures Markets*, 12, 519-538.
- Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, 36, 394-419.

- Massimb, M. N., & Phelps, B. D. (1994). Electronic Trading, Market Structure and Liquidity. *Financial Analysts Journal*, 50, 39-50.
- McInish, T., & Wood, R. (1992). An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks. *Journal of Finance*, 47, 753-763.
- Merton, R. C. (1973). The Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science*, 4, 141-183.
- Myers, R. J. (1991). Estimating Time-Varying Optimal Hedge Ratios on Commodity Futures Markets. *The Journal of Futures Markets*, 11, 39-53.
- Newey, W., & West, K. (1987). A Simple Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 51, 76-89.
- Pearl, J. (2000). *Causality*. Cambridge, MA: Cambridge University Press.
- Peck, A. E. (1975). Hedging and Income Stability: Concepts, Implications, and an Example. *American Journal of Agricultural Economics*, 57, 410-419.
- Petersen, M. A., & Fialkowski, D. (1994). Posted Versus Effective Spreads: Good Prices or Bad Quotes? *Journal of Financial Economics*, 35, 269-292.
- Pirrong, C. (1996). Market Liquidity and Depth on Computerized and Open Outcry Systems: A Comparison of DTB and LIFFE Bund Contracts. *The Journal of Futures Markets*, 5, 519-543.
- Reichenbach, H. (1956). *The Direction of Time*. Berkeley, CA: University of California Press.
- Rockwell, C. S. (1967). Normal Backwardation, Forecasting, and the Returns to Commodity Futures Traders. *Food Institute Research Studies*, 7, 107-130.

- Roll, R. (1984). A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *Journal of Finance*, 23, 1127-1139.
- Ross, S. A. (1978). A Simple Approach to the Valuation of Risky Streams. *The Journal of Business*, 51, 453-475.
- Samuelson, P. (1965). Proof That Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review*, 6, 41-49.
- Schwartz, E. S. (1997). The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. *Journal of Finance*, 52, 923-974.
- Schwarz, G. (1978). Estimating the Dimensions of a Model. *Annals of Statistics*, 6, 461-464.
- Shalen, C. T. (1993). Volume, Volatility, and the Dispersion of Beliefs. *Review of Financial Studies*, 6, 405-434.
- Shyy, G., Vijayraghavan, V., & Scott-Quinn, B. (1996). A Further Investigation of the Lead-Lag Relationship Between the Cash Market and Stock Index Futures Market With the Use of Bid-Ask Quotes: The Case of France. *Journal of Futures Markets*, 16, 405-420.
- Silber, W. (1984). Marketmaker Behavior in an Auction Market: An Analysis of Scalpers in Futures Markets. *Journal of Finance*, 4, 937-953.
- Smith, T., & Whaley, R. E. (1994). Estimating the Effective Bid/Ask Spread from Time and Sales Data. *Journal of Futures Markets*, 14, 437-456.
- Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, Prediction, and Search*. Cambridge, MA: The MIT Press.

- Stein, J. C. (1987). Informational Externalities and Welfare-Reducing Speculation. *Journal of Political Economy*, 95, 1123-1145.
- Stock, J. H. (1987). Asymptotic Properties of Least-Squares of Co-integrating Vectors. *Econometrica*, 55, 1035-1059.
- Stoll, H. (1978). The Pricing of Security Dealer Services: An Empirical Study of NASDAQ Stocks. *Journal of Finance*, 33, 1153-1172.
- Subrahmanyam, A. (1991). A Theory of Trading in Stock Index Futures. *Review of Financial Studies*, 4, 17-51.
- Swanson, N. R., & Granger, C. W. J. (1997). Impulse Response Functions Based on a Causal Approach to Residual Orthogonalization in Vector Autoregressions. *Journal of the American Statistical Association*, 92, 357-367.
- Tauchen, G. E., & Pitts, M. (1983). The Price Variability-Volume Relationship on Speculative Markets. *Econometrica*, 51, 485-506.
- Telser, L. (1958). Futures Trading and the Storage of Cotton and Wheat. *Journal of Political Economy*, 66, 233-255.
- Thompson, S., Eales, J. S., & Siebold, D. (1993). Comparison of Liquidity Costs Between the Kansas City and Chicago Wheat Futures Contracts. *Journal of Agriculture and Resource Economic*, 18, 185-197.
- Thompson, S. R., & Waller, M. (1988). Determinants of Liquidity Costs in Commodity Futures Markets. *Review of Futures Markets*, 7, 110-126.
- Tse, Y. (1999). Market Microstructure of FT-SE 100 Index Futures: An Intraday Empirical Analysis, *The Journal of Futures Markets*, 19, 31-58.

- Tse, Y. & Zobotina, T. V. (2001). Transaction Costs and Market Quality: Open Outcry Versus Electronic Trading, *The Journal of Futures Markets*, 21, 713-735.
- Venkataraman, K. (2001). Automated Versus Floor Trading: An Analysis of Execution Costs on the Paris and New York Exchanges. *Journal of Finance*, 56, 1445-1485.
- Wang, C. (2002). The Effect of Net Positions By Trader Type on Volatility in Foreign Currency Futures Markets. *Journal of Futures Markets*, 22, 427-450.
- Wang, H. K. W., Yau, J., & Baptiste, T. (1997). Trading Volume and Transaction Costs in Futures Markets. *Journal of Futures Markets*, 17, 757-780.
- Wang, J. (1999). Asymmetric Information and the Bid-Ask Spread: An Empirical Comparison Between Automated Order Execution and Open Outcry Auction. *Journal of International Financial Markets, Institutions, and Money*, 9, 115-128.
- White, H. (1980). A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrica*, 48, 817-838.
- Working, H. (1949). The Theory of Price of Storage. *American Economic Review*, 39, 1254-1262.

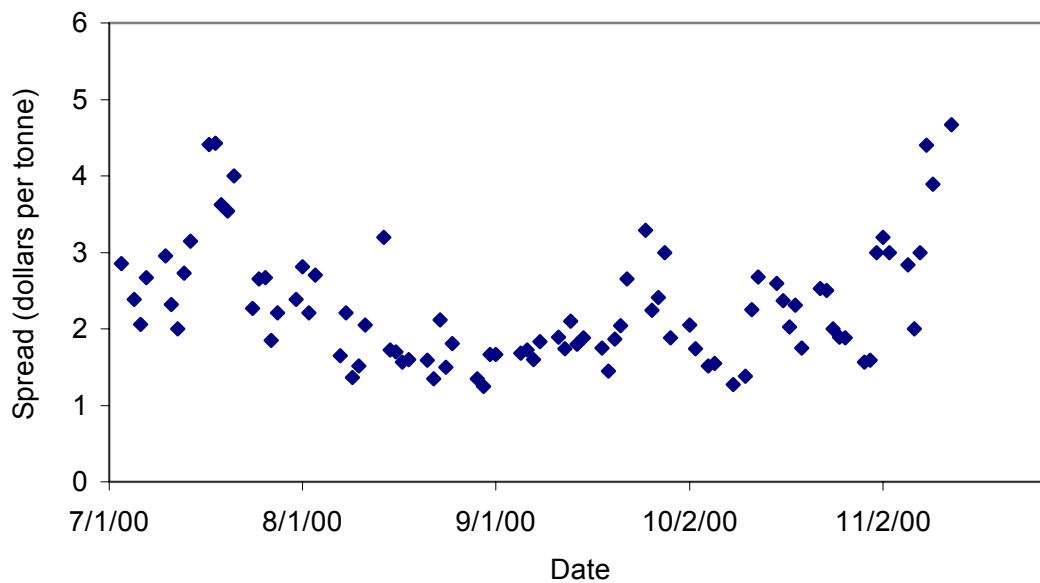
APPENDIX A**FIGURES**

Figure 2.1 Daily average bid-ask spread for November 2000 coffee futures (dollars per tonne)

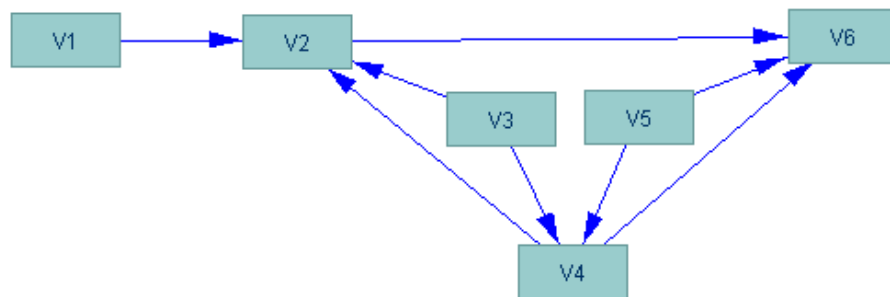


Figure 3.1 Directed acyclic graph G

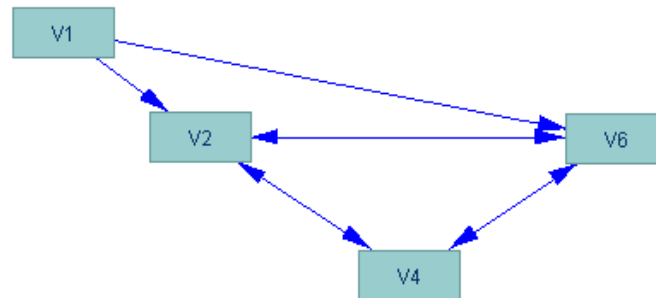


Figure 3.2 Inducing path graph G' over $O = \{V_1, V_2, V_4, V_6\}$ associated with directed acyclic graph G

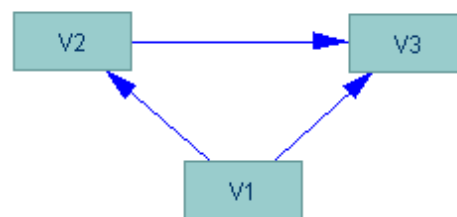


Figure 3.3 Directed acyclic graph H

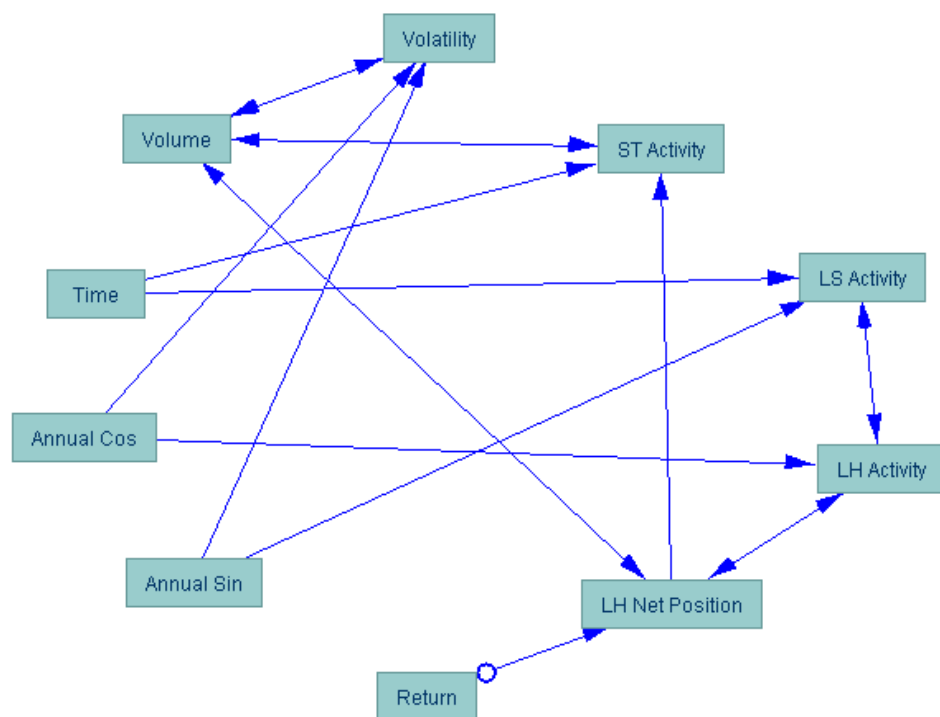


Figure 3.4 Partially-oriented inducing path graph for Corn

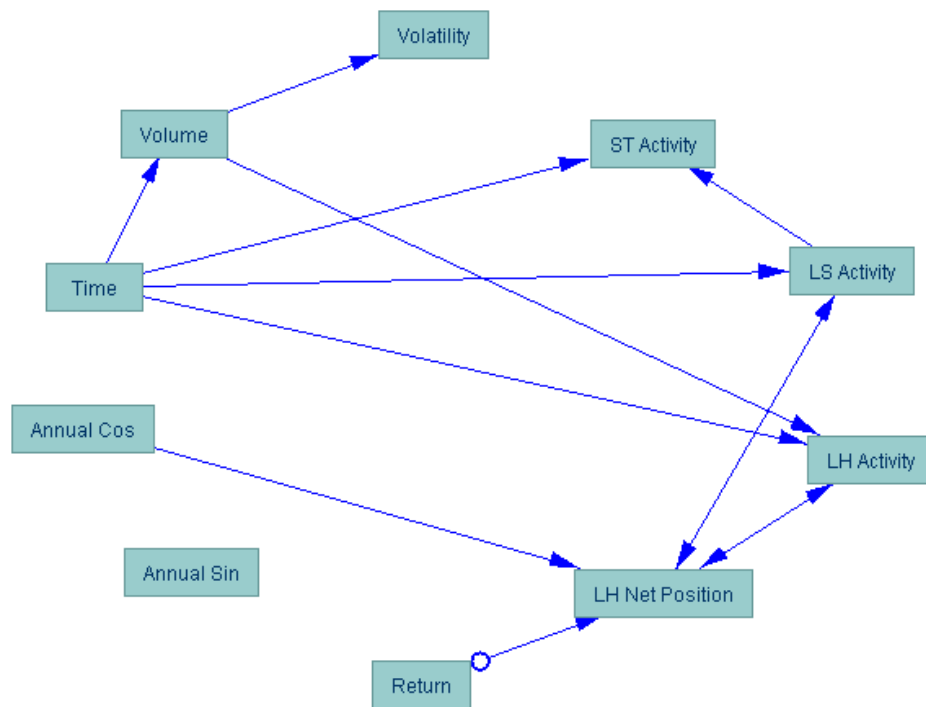


Figure 3.5 Partially-oriented inducing path graph for Crude Oil

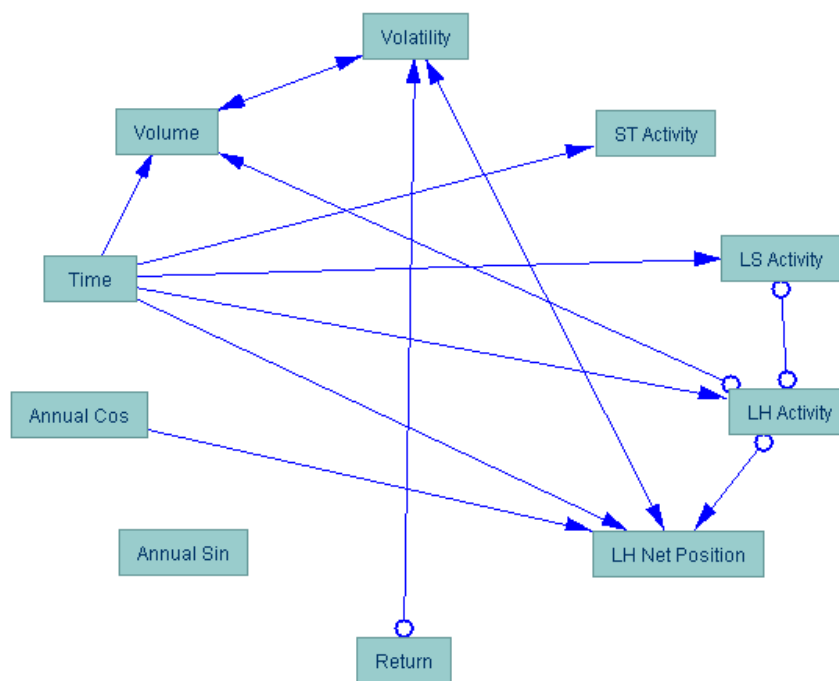


Figure 3.6 Partially-oriented inducing path graph for Eurodollars

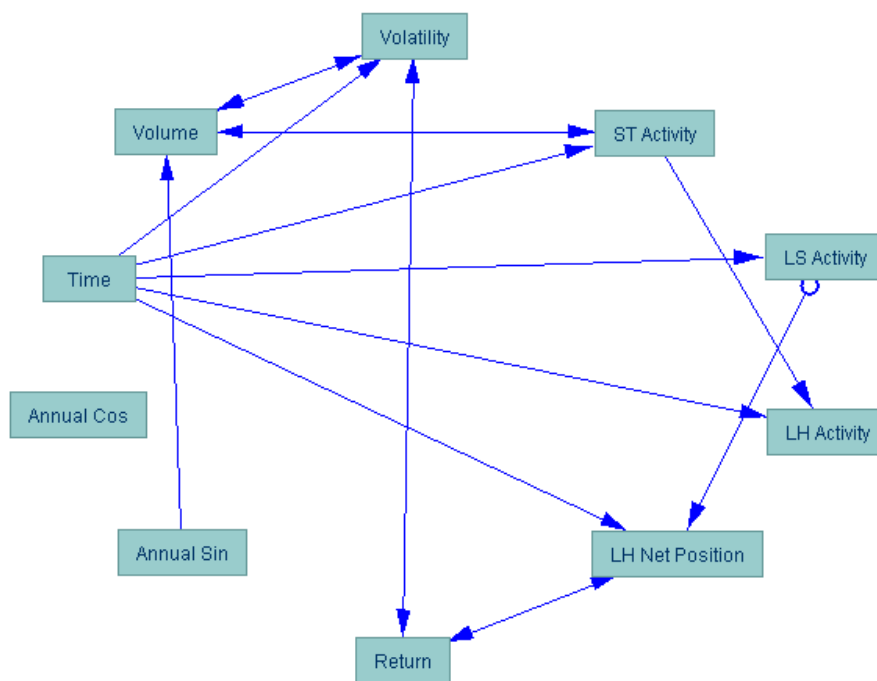


Figure 3.7 Partially-oriented inducing path graph for Gold

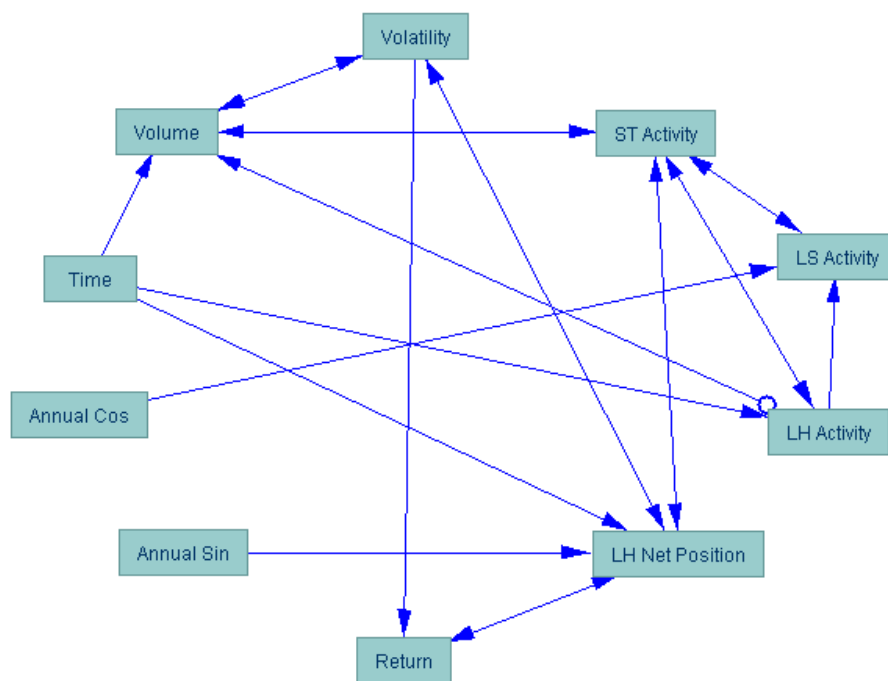


Figure 3.8 Partially-oriented inducing path graph for Japanese Yen

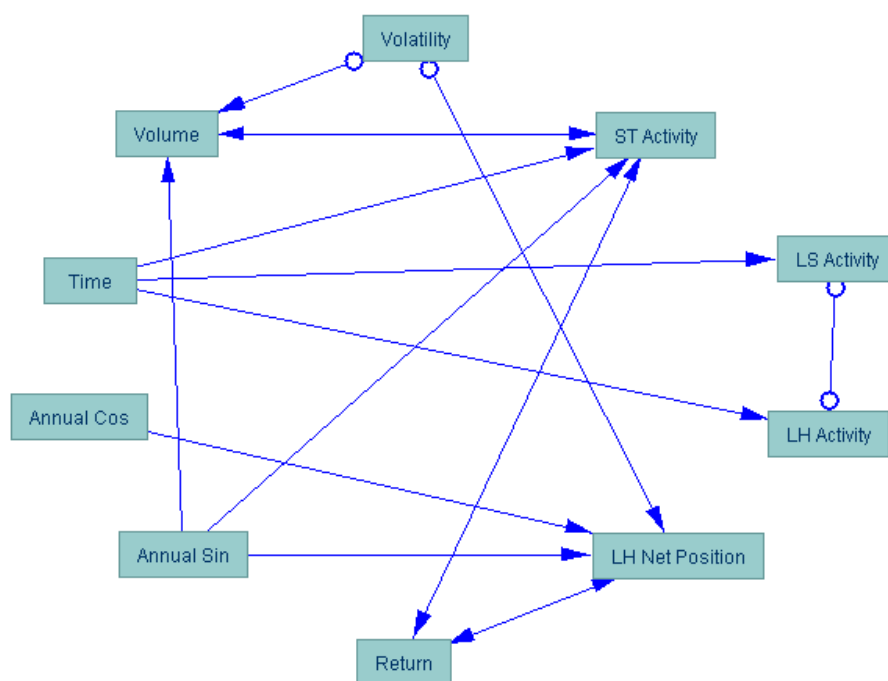


Figure 3.9 Partially-oriented inducing path graph for Coffee

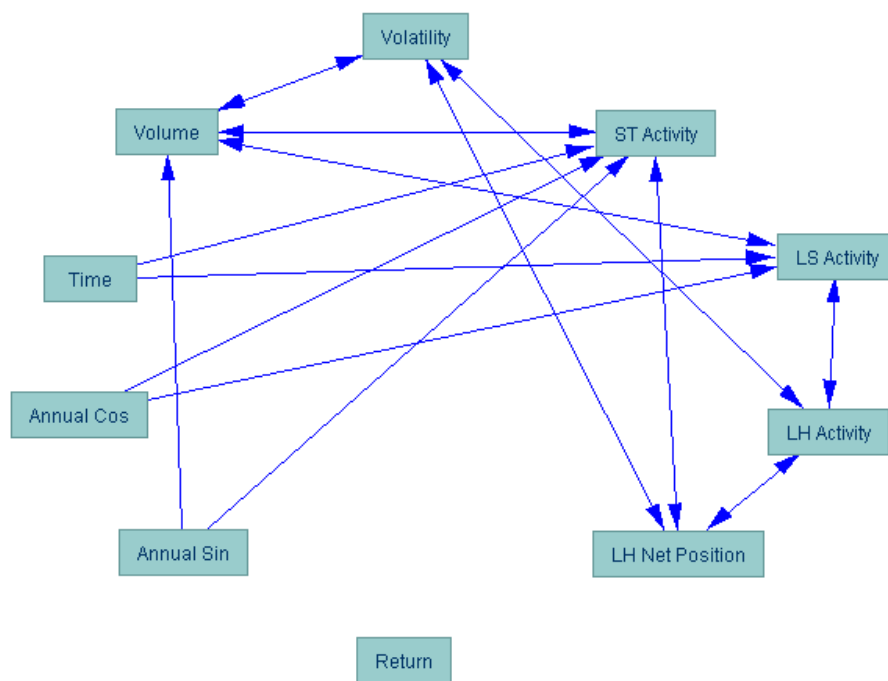


Figure 3.10 Partially-oriented inducing path graph for Live Cattle

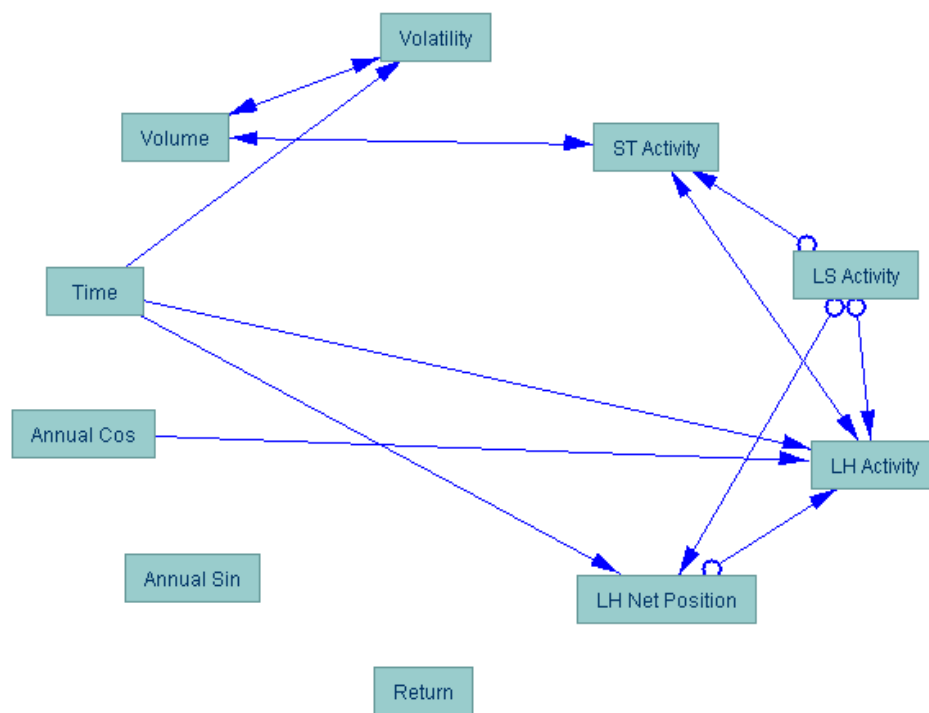


Figure 3.11 Partially-oriented inducing path graph for S&P 500

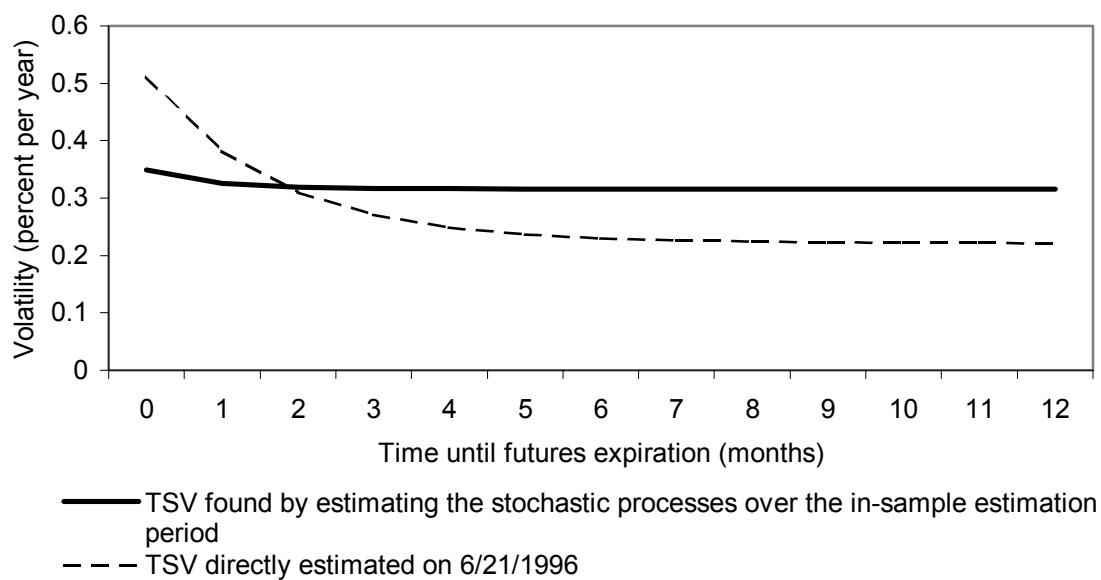


Figure 4.1 The term structure of volatility (TSV) of crude oil in the Gibson-Schwartz model using different parameter estimation techniques

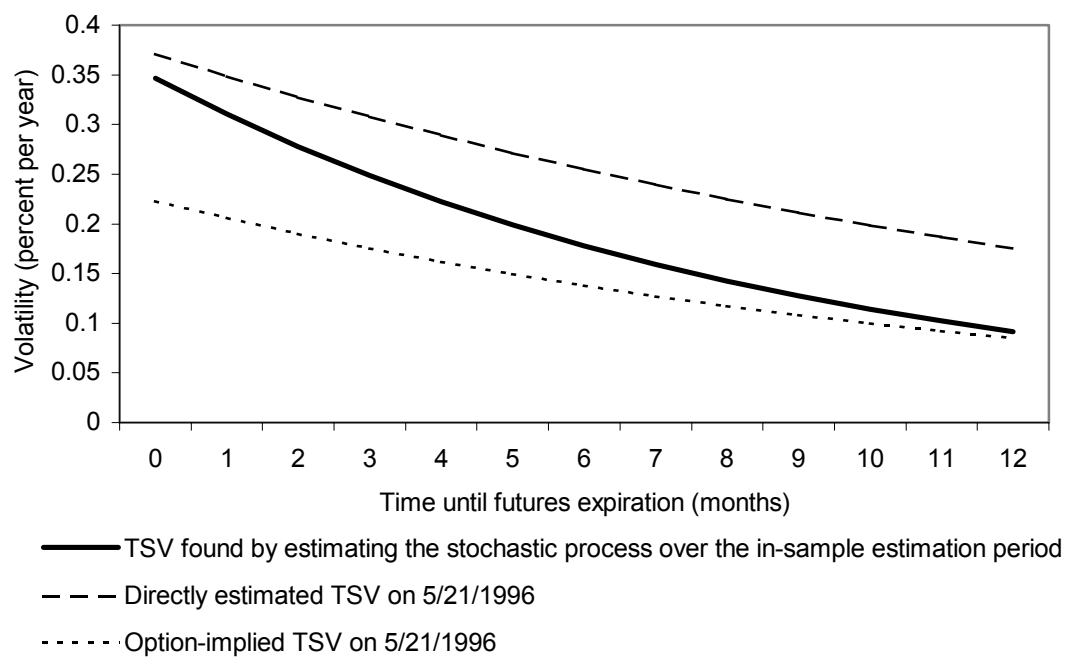


Figure 4.2 The term structure of volatility (TSV) of crude oil in the Schwartz 1997 model using different parameter estimation and inference techniques

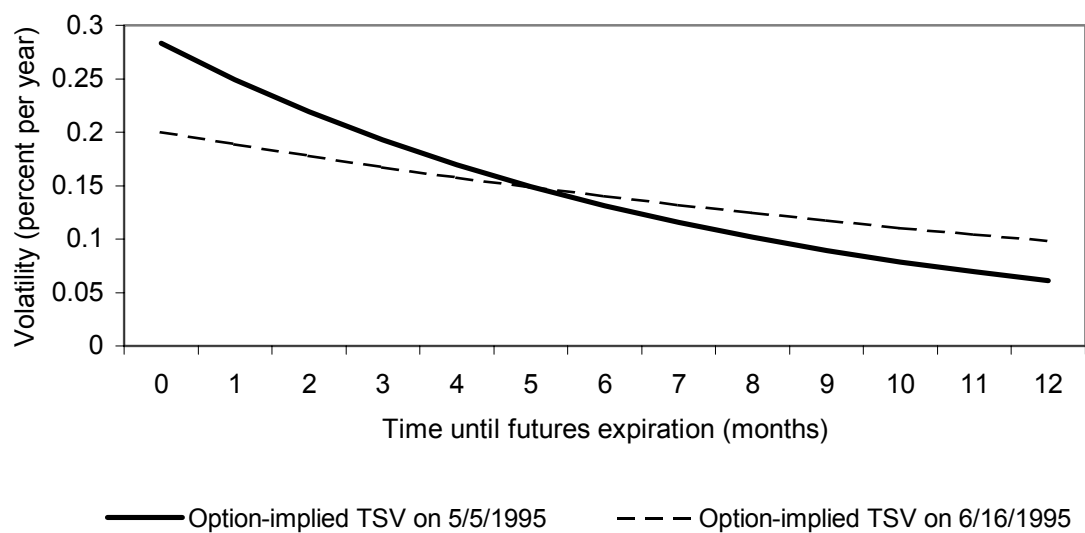


Figure 4.3 The option-implied term structure of volatility (TSV) of crude oil in the Schwartz 1997 model observed on two dates

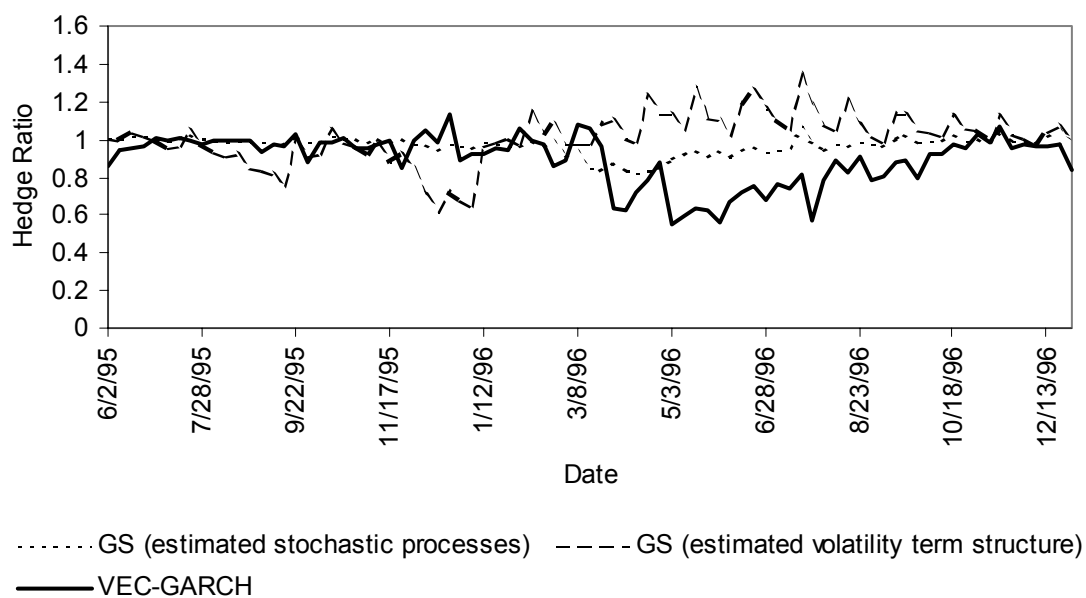


Figure 4.4 Partial time paths of the in-sample optimal hedge ratios generated by the VEC-GARCH and Gibson-Schwartz (GS) models

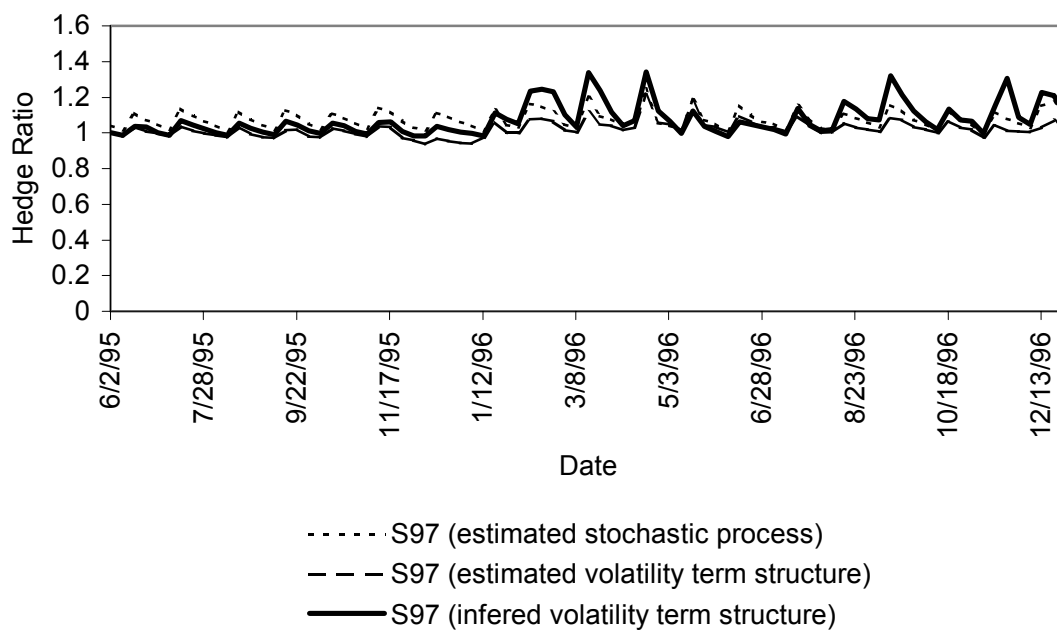


Figure 4.5 Partial time paths of the in-sample optimal hedge ratios generated by the Schwartz 1997 (S97) models

APPENDIX B

TABLES

Table 2.1 Example of LIFFE coffee futures data^a

<i>Date</i>	<i>Time</i>	<i>Delivery</i>	<i>Type</i>	<i>Volume</i>	<i>Price</i>
10/27/00	10:03:50	Nov-00	Bid	0	701
10/27/00	10:04:12	Nov-00	Bid	0	702
10/27/00	10:04:49	Nov-00	Ask	0	702
10/27/00	10:04:50	Nov-00	Bid	0	701
10/27/00	10:04:51	Nov-00	Trd	3	702
10/27/00	10:05:16	Nov-00	Ask	0	703
10/27/00	10:05:31	Nov-00	Trd	5	701
10/27/00	10:05:45	Nov-00	Trd	5	701
10/27/00	10:07:09	Nov-00	Trd	20	703
10/27/00	10:08:18	Nov-00	Bid	0	702
10/27/00	10:11:12	Nov-00	Trd	20	702
10/27/00	10:11:24	Nov-00	Trd	1	703
10/27/00	10:18:15	Nov-00	Ask	0	702
10/27/00	10:18:16	Nov-00	Bid	0	701
10/27/00	10:19:37	Nov-00	Trd	1	702
10/27/00	10:19:38	Nov-00	Trd	1	702
10/27/00	10:19:41	Nov-00	Trd	1	701

^a Source: London International Financial Futures and Options Exchange (LIFFE). “Type” refers to type of price observation. “Trd” denotes a trade observation.

Table 2.2 Deviations of intra-day average spreads from overall daily average spreads

<i>Cocoa</i>					<i>Coffee</i>				
Intra-day Period	Mean	Standard deviation	t-stat	p-value	Intra-day Period	Mean	Standard deviation	t-stat	p-value
1	-0.238	0.658	-0.362	0.718	1	-0.226	0.470	-0.482	0.631
2	0.083	0.531	0.156	0.877	2	0.152	0.807	0.188	0.851
3	0.239	0.499	0.478	0.633	3	0.215	0.665	0.323	0.747
4	0.012	0.380	0.030	0.976	4	-0.002	0.428	-0.005	0.996
5	0.016	0.485	0.032	0.974	5	0.182	0.603	0.301	0.764
6	0.121	0.366	0.330	0.742	6	0.150	0.627	0.240	0.811

Table 2.3 Correlations of daily average spreads and estimates of daily average spreads^b

<i>Cocoa</i>	<i>RM</i>	<i>CDP</i>	<i>TWM</i>	<i>CFTC</i>	<i>SW</i>	<i>Spread</i>
<i>RM</i>	1.00	0.71	0.49	0.50	0.46	0.32
<i>CDP</i>		1.00	0.57	0.60	0.51	0.20
<i>TWM</i>			1.00	0.85	0.96	0.60
<i>CFTC</i>				1.00	0.84	0.47
<i>SW</i>					1.00	0.59
<i>Coffee</i>	<i>RM</i>	<i>CDP</i>	<i>TWM</i>	<i>CFTC</i>	<i>SW</i>	<i>Spread</i>
<i>RM</i>	1.00	0.70	0.41	0.43	0.20	0.12
<i>CDP</i>		1.00	0.55	0.63	0.23	0.24
<i>TWM</i>			1.00	0.93	0.93	0.85
<i>CFTC</i>				1.00	0.86	0.82
<i>SW</i>					1.00	0.80

^b *RM*: Roll's measure; *CDP*: Chu, Ding and Pyun estimator; *TWM*: Thompson-Waller measure; *CFTC*: Commodity Futures Trading Commission estimator; *SW*: Smith and Whaley estimator.

Table 2.4 Performance of estimators by commodity^c

	<i>RM</i>	<i>CDP</i>	<i>TWM</i>	<i>CFTC</i>	<i>SW</i>	<i>RM</i>	<i>CDP</i>	<i>TWM</i>	<i>CFTC</i>	<i>SW</i>
<i>Cocoa</i>	<i>Pounds per tonne</i>					<i>Pounds per contract</i>				
Mean error	-0.77	-0.52	-0.18	-0.17	-0.21	-7.74	-5.19	-1.84	-1.65	-2.14
Mean squared error	0.73	0.52	0.08	0.10	0.09	7.26	5.25	0.84	0.95	0.91
Root mean squared error	0.85	0.72	0.29	0.31	0.30	8.52	7.24	2.89	3.08	3.02
Mean absolute error	0.78	0.62	0.23	0.23	0.24	7.82	6.18	2.26	2.32	2.39
Mean absolute percent error	51.72	40.80	14.15	14.45	14.84					
Total number of observations	111	100	149	149	148					
Serial correlation errors	38	49	N/A	N/A	N/A					
<i>Coffee</i>	<i>Dollars per tonne</i>					<i>Dollars per contract</i>				
Mean error	-1.02	-0.76	-0.47	-0.44	-0.55	-5.10	-3.82	-2.34	-2.19	-2.74
Mean squared error	1.33	0.91	0.31	0.30	0.44	6.67	4.53	1.57	1.52	2.19
Root mean squared error	1.15	0.95	0.56	0.55	0.66	5.77	4.76	2.80	2.76	3.31
Mean absolute error	1.02	0.78	0.48	0.45	0.55	5.10	3.91	2.40	2.27	2.75
Mean absolute percent error	51.00	39.03	22.37	20.97	25.03					
Total number of observations	123	117	143	143	137					
Serial correlation errors	20	26	N/A	N/A	N/A					

^c*RM*: Roll's measure; *CDP*: Chu, Ding and Pyun estimator; *TWM*: Thompson-Waller measure; *CFTC*: Commodity Futures Trading Commission estimator; *SW*: Smith and Whaley estimator.

Table 2.5 Coefficient estimates and p -value for differences in bias and variance components for each pair of bid-ask spread estimators^d

	β_0				β_1			
<i>Cocoa</i>	<i>CDP</i>	<i>TWM</i>	<i>CFTC</i>	<i>SW</i>	<i>CDP</i>	<i>TWM</i>	<i>CFTC</i>	<i>SW</i>
<i>RM</i>	0.230 (0.000)	0.586 (0.000)	0.624 (0.000)	0.563 (0.000)	0.205 (0.000)	-0.319 (0.000)	-0.219 (0.000)	-0.358 (0.000)
<i>CDP</i>		0.324 (0.000)	0.378 (0.000)	0.300 (0.000)		-0.475 (0.000)	-0.397 (0.000)	-0.509 (0.000)
<i>TWM</i>			0.019 (0.084)	-0.032 (0.000)			0.084 (0.001)	-0.020 (0.137)
<i>CFTC</i>				-0.052 (0.000)				-0.103 (0.000)
<i>Coffee</i>	<i>CDP</i>	<i>TWM</i>	<i>CFTC</i>	<i>SW</i>	<i>CDP</i>	<i>TWM</i>	<i>CFTC</i>	<i>SW</i>
<i>RM</i>	0.219 (0.000)	0.566 (0.000)	0.615 (0.000)	0.522 (0.000)	0.077 (0.022)	-0.340 (0.000)	-0.301 (0.000)	-0.275 (0.000)
<i>CDP</i>		0.316 (0.000)	0.369 (0.000)	0.290 (0.000)		-0.385 (0.000)	-0.329 (0.000)	-0.338 (0.000)
<i>TWM</i>			0.031 (0.043)	-0.079 (0.000)			0.048 (0.054)	0.106 (0.000)
<i>CFTC</i>				-0.096 (0.000)				0.076 (0.014)

^d *RM*: Roll's measure; *CDP*: Chu, Ding and Pyun estimator; *TWM*: Thompson-Waller measure; *CFTC*: Commodity Futures Trading Commission estimator; *SW*: Smith and Whaley estimator. $\beta_0 > 0$ implies that the bias of the estimator in the row is greater than the bias of the estimator in the column. $\beta_0 < 0$ implies the opposite. $\beta_1 > 0$ implies that the variance of the estimator in the row is greater than the variance of the estimator in the column. $\beta_1 < 0$ implies the opposite. P – values close to zero suggest that the bias and or/variance of two estimators is statistically different

Table 2.6 P-values for encompassing tests^e

<i>Cocoa</i>	<i>RM</i>	<i>CDP</i>	<i>TWM</i>	<i>CFTC</i>	<i>SW</i>
<i>RM</i>		0.000	0.000	0.000	0.000
<i>CDP</i>	0.583		0.000	0.000	0.000
<i>TWM</i>	0.000	0.000		0.217	0.429
<i>CFTC</i>	0.000	0.000	0.000		0.001
<i>SW</i>	0.000	0.000	0.000	0.007	
<i>Coffee</i>	<i>RM</i>	<i>CDP</i>	<i>TWM</i>	<i>CFTC</i>	<i>SW</i>
<i>RM</i>		0.000	0.000	0.000	0.000
<i>CDP</i>	0.059		0.000	0.000	0.000
<i>TWM</i>	0.000	0.000		0.010	0.000
<i>CFTC</i>	0.000	0.000	0.144		0.011
<i>SW</i>	0.000	0.000	0.000	0.000	

^e *RM*: Roll's measure; *CDP*: Chu, Ding and Pyun estimator; *TWM*: Thompson-Waller measure; *CFTC*: Commodity Futures Trading Commission estimator; *SW*: Smith and Whaley estimator. *P*-values are for the test of H_0 : the estimator in a row encompasses the estimator in a column. A *p*-value close to zero suggests that the estimator in a particular row does not encompass an estimator in a particular column.

Table 2.7 Determinants of daily average spreads regression results^f

	<i>Cocoa</i>		<i>Coffee</i>	
	Coefficient estimate	p-value	Coefficient estimate	p-value
<i>Constant</i>	-6.726	0.000	-2.105	0.050
<i>D_e</i>	0.637	0.090	0.678	0.020
<i>Sqr(volume)</i>	0.000	0.865	-0.006	0.109
<i>Sqr(variance)</i>	0.071	0.450	0.068	0.002
<i>D_eSqr(volume)</i>	-0.018	0.015	-0.003	0.719
<i>D_eSqr(variance)</i>	0.469	0.036	0.083	0.078
<i>Sqr(price)</i>	0.329	0.000	0.142	0.000
<i>R</i> ²	0.715		0.336	

^f *D_e* is a dummy variable that is zero for an open-outcry observation and one for an electronic observation, *volume* is the total volume of futures traded, *variance* is the variance of spread midpoints, and *price* is the average spread midpoint for a day.

Table 3.1 Results of augmented Dickey-Fuller tests^g

Commodity	<i>Nearby</i>	<i>Return</i>	<i>Volatility</i>	<i>Volume</i>	<i>LH Activity</i>	<i>LS Activity</i>	<i>ST Activity</i>	<i>LH Net Position</i>
Corn	-2.73 (1)	-19.59 (0)	-6.53 (2)	-11.49 (0)	-3.70 (1)	-4.10 (0)	-3.42 (1)	-3.94 (1)
Crude Oil	-2.03 (1)	-21.66 (0)	-15.53 (0)	-6.65 (3)	-2.30 (13)	-3.54 (0)	-1.76 (9)	-4.78 (1)
Eurodollars	-0.24 (0)	-19.07 (0)	-6.67 (2)	-5.44 (3)	-2.60 (13)	-2.81 (16)	-3.79 (1)	-4.01 (0)
Gold	0.02 (0)	-19.20 (0)	-14.17 (0)	-12.91 (0)	-3.47 (0)	-4.23 (1)	-3.38 (2)	-4.33 (1)
Japanese Yen	-2.25 (0)	-18.49 (0)	-5.97 (3)	-2.08 (12)	-4.17 (13)	-5.48 (0)	-7.73 (0)	-5.81 (1)
Coffee	-2.69 (3)	-21.39 (0)	-9.50 (1)	-14.11 (0)	-3.72 (0)	-4.55 (0)	-4.89 (1)	-5.90 (1)
Live Cattle	-3.82 (0)	-17.95 (1)	-6.78 (2)	-14.25 (0)	-1.96 (0)	-4.96 (1)	-4.77 (10)	-3.30 (1)
S&P 500	-0.24 (1)	-24.26 (0)	-8.91 (1)	-1.71 (12)	-1.97 (14)	-3.18 (13)	-1.92 (14)	-3.45 (0)

^g The null hypothesis is that the series listed in the row and column intersection has a unit root. We reject this hypothesis if the ADF test statistic is less than the critical value -3.13 (10%) given in Fuller (1976). Both an intercept and a time trend were included in the tests. The optimal lag length given in parenthesis was chosen using the Schwarz (1978) information criterion.

Table 3.2 Conditioning sets that result in vanishing correlations^h

Market	Edge between <i>LS Activity</i> and <i>Volatility</i>	Edge between <i>ST Activity</i> and <i>Volatility</i>	Edge between <i>LS Activity</i> and <i>Volume</i>	Edge between <i>ST Activity</i> and <i>Volume</i>
Corn	<i>Annual Sin</i>	Empty Set	Empty Set	(None)
Crude Oil	<i>LH Activity</i>	<i>Volume</i>	<i>ST Activity, LH Net Position, LH Time</i>	<i>LS Activity, Time</i>
Eurodollars	<i>Return</i>	<i>Volume</i>	<i>LH Activity</i>	<i>LS Activity</i>
Gold	<i>Volume, Time</i>	Empty Set	<i>Volatility</i>	(None)
Japanese Yen	Empty Set	Empty Set	<i>LH Activity</i>	(None)
Coffee	Empty Set	Empty Set	<i>LH Activity</i>	(None)
Live Cattle	<i>LH Net Position</i>	Empty Set	(None)	(None)
S&P 500	<i>Time</i>	<i>Time</i>	<i>Volatility</i>	(None)

^h For the market listed in a row, the correlation between the pair of variables listed in the column is not significantly different from zero, conditional on the variables given in the row and column intersection. “(None)” indicates that no set of variables is found that results in a correlation not significantly different from zero. “Empty Set” indicates that the unconditional correlation between the two variables is not significantly different from zero. *Return* is the log change in the price of the nearby futures contract, *Volume* is the total volume of trade, *Volatility* is the log difference between the high and low nearby futures prices for a week, *LH Net Position* is net futures position of large hedgers, *LH Activity*, *LS Activity*, and *ST Activity* are the total number of open futures positions of large hedger, large speculators, and small traders, respectively. *Time* is a linear time trend, and *Annual Sin* is an annual seasonal harmonic variable.

Table 4.1 Results from augmented Dickey-Fuller tests on price dataⁱ

Series	K	θ_1
Spot Price	0	-2.907
Spot Price Changes	0	-28.819
Nearby Future Price	2	-2.973
Nearby Futures Price Changes	1	-16.474
ECT	3	-11.458

ⁱ Tests for the presence of unit roots, using an intercept but no time trend. The critical value -3.43 (1%) is given in Fuller (1976). The optimal lag length (K) was chosen using the Schwarz (1978) information criterion.

Table 4.2 Parameter estimates and residual diagnostics for the univariate GARCH(1,1) models^j

	<i>DSPOT</i>	<i>NEARD</i>
μ	-0.006 (0.027)	0.057 (0.028)
ω	0.051 (0.013)	0.089 (0.019)
α	0.184 (0.037)	0.200 (0.038)
β	0.764 (0.042)	0.679 (0.052)
Log-likelihood	-197.416	-182.392
m^3	-0.330	-0.069
m^4	2.180	2.504
$Q(12)$	17.541 (0.130)	17.249 (0.140)
$Q^2(12)$	6.745 (0.874)	9.581 (0.653)

^jThe model is given by:

$$x_t = \mu + \varepsilon_t$$

$$\varepsilon_t \mid \Omega_{t-1} \sim N(0, h_t^2)$$

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}$$

The numbers in parenthesis beside the parameter estimates are asymptotic standard errors. m^3 and m^4 are the sample skewness and sample kurtosis, respectively, of the standardized residuals. $Q(12)$ and $Q^2(12)$ denote Ljung-Box test statistics for 12th-order autocorrelation in the standardized and squared standardized residuals, respectively, with the numbers in parenthesis being the associated p-values.

Table 4.3 Parameter estimates and residual diagnostics for the multivariate GARCH(1,1) model^k

α_1	-0.572	(0.102)
α_2	0.299	(0.106)
W_1	0.078	(0.006)
W_2	0.069	(0.001)
W_3	0.070	(0.005)
A_{11}	0.121	(0.011)
A_{22}	0.093	(0.009)
A_{22}	0.096	(0.011)
B_{11}	0.761	(0.014)
B_{22}	0.785	(0.011)
B_{33}	0.786	(0.018)
Log-likelihood	412.628	
<i>DSPOT equation</i>		
m^3	-0.264	
m^4	2.417	
$Q(12)$	17.157	(0.144)
$Q^2(12)$	11.972	(0.448)
<i>NEARD equation</i>		
m^3	-0.150	
m^4	2.795	
$Q(12)$	17.067	(0.147)
$Q^2(12)$	15.493	(0.216)

^kThe model is given by:

$$\Delta y_t = \alpha ECT_{t-1} + \varepsilon_t; \quad \Delta y_t = (DSPOT_t, NEARD_t)^T$$

$$\varepsilon_t | \Omega_{t-1} \sim N(0, H_t)$$

$$\text{vech}(H_t) = W + A \text{vech}(\varepsilon_{t-1} \varepsilon_{t-1}^T) + B \text{vech}(H_{t-1})$$

The numbers in parenthesis beside the parameter estimates are asymptotic standard errors. m^3 and m^4 are the sample skewness and sample kurtosis, respectively, of the standardized residuals. $Q(12)$ and $Q^2(12)$ denote Ljung-Box test statistics for 12th-order autocorrelation in the standardized and squared standardized residuals, respectively, with the numbers in parenthesis being the associated p-values.

Table 4.4 In-sample hedging effectiveness¹

	Average CEI
Unhedged	-5.648E+09
S97 (estimated SDE)	-1.018E+09
S97 (estimated TSV)	-9.537E+08
S97 (inferred TSV)	-1.058E+09
GS (estimated SDEs)	-9.498E+08
GS (estimated TSV)	-1.143E+09
VEC-GARCH	-8.380E+08

¹CEI is certainty equivalent income.

Table 4.5 Out-of-sample hedging effectiveness^m

	Average CEI
Unhedged	-9.370E+09
S97 (estimated SDE)	-2.908E+09
S97 (estimated TSV)	-2.723E+09
S97 (inferred TSV)	-2.989E+09
GS (estimated SDEs)	-2.773E+09
GS (estimated TSV)	-2.954E+09
VEC-GARCH	-2.281E+09

^m CEI is certainty equivalent income.

VITA

Henry L. Bryant IV was born in Houston, Texas. He graduated with a B. S. in Business Administration from the University of Nevada, Las Vegas, in 1991. He opened and ran a small grocery store from 1991 through 1995. He has been an active speculator in derivatives markets since 1994. He began pursuit of a Ph.D. in Agricultural Economics at Texas A&M University in 1998, receiving his degree in August 2003. Correspondence can reach Mr. Bryant c/o Henry L. Bryant III, P.O. Box 187, Dublin, TX 76446.